
Understanding Periodontal Research

VISIT...

LANZAROTE
Caliente.COM

Alexandrina L. Dumitrescu

Understanding Periodontal Research

Alexandrina L. Dumitrescu
University of Tromsø
Institute of Clinical Dentistry
Department of Periodontology
Tromsø
Norway

Private practice
Bucharest
Romania

ISBN 978-3-642-28922-4 ISBN 978-3-642-28923-1 (eBook)
DOI 10.1007/978-3-642-28923-1
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012945294

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*The noblest pleasure is the joy of
understanding.*

Leonardo da Vinci

Preface

The concept of evidence-based medicine has become a commonplace phrase, representing the hallmark of excellence in clinical practice. The foundation for evidence-based practice was laid by David Sackett (1997) who has defined it as “integrating individual clinical experience with the best available external clinical evidence from systematic research.” A Medline search of “evidence-based medicine” in 1993 revealed 6 citations, and now in 2011, there are 73,519, while search of “evidence-based dentistry” revealed now 1129 citations and none 20 years ago. But does this dramatic growth and awareness represent better research and clinical practice?

Dental clinicians must understand the importance of the research question, study design, and outcomes to apply the best available evidence to patient care. The most compelling and growing component of evidence-based dentistry is the patient’s empowerment in the treatment decision-making process. Evidence-based decision aids are being developed and evaluated to help patients make more informed and personalized choices about treatment options. The true philosophy of evidence-based dentistry, however, is not for research to supplant individual clinical experience and the patients’ informed preference, but to integrate them with the best available research.

This book is meant to introduce dental practitioners to the concepts of evidence-based medicine by offering them a bridge from science to clinical practice, and to propose young researchers a guide to understand the periodontal literature published in the last 30 years. I just want to give credit and honor those dental researchers and scientists who labored in this field, giving their devotion to expanding critical knowledge, which will benefit patients and clinicians alike.

A.L.D.

Contents

1 Evidence-Based Periodontology	1
1.1 What Is Evidence-Based Medicine?	1
1.2 What Is Evidence-Based Dentistry?	6
1.3 What Is Evidence-Based Periodontology?	6
1.4 Evidence-Based Periodontology Versus Traditional Periodontology	6
References	6
2 Ethical Considerations in Periodontal Research	9
2.1 Protection of Human Study Participants.	9
2.2 Responsibilities of Investigators	19
2.3 Ethical Issues Specific to Certain Types of Research	22
References	27
3 Overview of Clinical Research Design Architecture	31
3.1 Developing a Research Question	31
3.2 Classification of Types of Clinical Research	32
3.3 Observational Clinical Research.	34
3.4 Experimental Studies	64
3.5 Small Clinical Trials	78
3.6 Sampling Methods	79
3.7 Measurement of Outcomes	81
3.8 Are the Findings Valid?	82
3.9 Assessing the Quality of Published Observational Research	82
References	83
4 Randomized Controlled Clinical Trials	89
4.1 Clinical Trial Design	89
4.2 Clinical Trial Implementation	115
4.3 Consolidated Standards of Reporting Trials (Consort) Guidelines	123
4.4 Special Types of RCTs: Comparative RCTs.	133
4.5 Special Types of RCTs: Cluster (Group)-Randomized Trials	137
References	138

5 Bias in Periodontal Research	145
5.1 Bias Before a Trial Begins	145
5.2 Bias During Trial Execution	151
5.3 Bias After a Trial Ends	153
5.4 Overarching Biases	154
5.5 The Effect of Bias on the Magnitude of Clinical Outcomes in Periodontology.	155
References	156
6 Measurement Reliability in Periodontal Research	159
6.1 Basic Terminology	159
6.2 Sources of Variability for the Measurement Reliability ..	159
6.3 Standardization Versus Calibration.	166
6.4 Reliability Measures	167
6.5 Inter- Versus Intra-examiner Variability	172
6.6 Intra-examiner Variability in Periodontal Studies.	173
References	180
7 Compliance Considerations in Clinical Trial Design	185
7.1 Compliance in Periodontal Antimicrobial Treatment. ...	185
7.2 Enhancing Adherence in Clinical Research	186
References	189
8 Performing Systematic Reviews and Meta-Analyses	191
8.1 Introduction.	191
8.2 Steps in Conducting a Systematic Review	195
8.3 Reporting of Systematic Reviews and Meta-analysis of Observational Studies	232
8.4 Reporting of Systematic Reviews and Meta-analysis of Randomized Controlled Trials	232
8.5 Assessing the Quality of Systematic Reviews	234
8.6 Pitfalls in Systematic Reviews	238
8.7 When Can a Meta-analysis Mislead?	242
8.8 Particular Types of Systematic Reviews and Meta-Analysis	242
8.9 Improving and Updating Reviews	246
References	250
9 Statistical Versus Clinical Significance in Periodontal Research	257
9.1 The Matter of Statistical Significance.	257
9.2 Definition of Clinical Significance	259
9.3 Methods and Criteria to Evaluate Clinical Relevance ...	260
9.4 The ‘Number Needed to Treat’ (NNT).	262
9.5 Clinical Significance of Periodontal Therapies.	266
References	268
10 Clinical Practice Guidelines in Periodontology	271
10.1 Benefits of Clinical Guidelines	271
10.2 Limits of Clinical Guidelines	272
10.3 Development of Clinical Practice Guidelines and Clinical Guidance Statements	274

10.4	Clinical Practice Guidelines Update Process	290
10.5	Clinical Guidelines That Support Common Dental Procedures	291
	References	294
11	Multilevel Modelling in Periodontal Research	297
11.1	Site-Based, Tooth-Based, and Subject-Based Periodontal Evaluation	297
11.2	Principles of Multilevel Modelling (MLM)	298
11.3	The MLwiN Software	300
11.4	The Application of Multilevel Modelling to Cross-Sectional Periodontal Data	300
11.5	The Application of Multilevel Modelling to Longitudinal Periodontal Data	302
11.6	The Application of Multilevel Modelling to Periodontal Treatment Response Data	305
	References	312
12	Contemporary Concepts in the Diagnosis of Periodontal Disease	315
12.1	Clinical Methods for Periodontal Diagnosis	315
12.2	Radiographic Assessment of Periodontal Disease	325
12.3	Newly Emerging Methods for Periodontal Diagnosis	330
12.4	Advances in Microbiologic Analysis	346
12.5	Biochemical Analysis as Part of Periodontal Diagnosis	370
	References	404
13	A Guide to Diagnostic Accuracy Studies in Periodontology	423
13.1	Types of Diagnostic Evaluation Studies	423
13.2	Used Terminology	424
13.3	The Differences Between Studies of Diagnosis and Therapy	426
13.4	Defining and Interpreting the Accuracy Diagnostic Test	427
13.5	Guide for Reporting Studies of Diagnostic Tests: The STARD Initiative	430
13.6	Checklists for Assessing Studies on Diagnostic Test Accuracy	432
13.7	Systematic Reviews and Meta-analysis of Diagnostic Studies	434
13.8	Application of GRADE: Making Evidence-Based Recommendations About Diagnostic Tests in Clinical Practice Guidelines	438
	References	440
14	Clinical Trials Principles for Evaluation of Antimicrobial Drugs in Periodontal Disease Treatment	441
14.1	Studies <i>In Vitro</i>	442
14.2	Studies <i>In Vivo</i>	443
14.3	Class or Trial Type	452
	References	453

15 The Hawthorne Effect	459
15.1 Terminology	459
15.2 The Placebo Effect	459
15.3 The Hawthorne Effect	460
References	462
16 Clinical Trial Principles for Evaluation of Products	
Used in the Management of Oral Malodour	465
16.1 Etiopathogenesis of Oral Malodour	465
16.2 Assessment Methods of Oral Malodour	467
16.3 Treatment of Oral Malodour	469
16.4 Guidelines for Investigating Subjects with Halitosis	470
16.5 The ADA Guidelines on Oral Malodour Products	473
References	475
17 Challenges in Periodontal Genetics Research	477
17.1 Periodontitis Is a Complex Disease	477
17.2 Polygenetic Model for Periodontitis	482
17.3 Epigenetics and Periodontal Disease	485
17.4 Challenges for Periodontitis Genetics Studies	487
17.5 Limitations of Genetic Association Studies	503
References	505
18 Quantification of Gingival, Periodontal, and Oral	
Hygiene Status in Periodontal Research Environment	507
18.1 Evaluation of Inflammatory Changes of the Gingiva	507
18.2 Indices for Measuring Periodontitis	525
18.3 Oral Hygiene Indices	540
18.4 Miscellaneous Indices	563
18.5 Treatment Needs Indices	564
References	568
19 Aspects of the Research Methodology for Periodontal	
Disease Assessment in Epidemiological Surveys	575
19.1 Introduction	575
19.2 Epidemiological Study Designs	575
19.3 Issues of Study Design	577
19.4 Definitions of Periodontitis and Methods That Have Been Used to Identify This Disease in Epidemiological Studies	589
19.5 Periodontal Disease Progression	619
19.6 Monitoring and Quality Control	621
19.7 Self-Reported Periodontal Disease	626
19.8 Using Administrative Data for Epidemiological Research	634
19.9 Risk Factors of Periodontal Disease	635
References	637

20 Assessing Confounding and Effect Modification of Periodontitis–Systemic Disease Associations	645
20.1 Confounding	645
20.2 Controlling for Confounding	649
20.3 Effect Modification	652
References	653
21 Animal Models in Periodontal Research	655
21.1 Animal Models Used in Periodontal Research	655
21.2 Selection of an Experimental Animal Model	655
21.3 Experimental Periodontal Defects	657
21.4 Nonhuman Primates	670
21.5 Dogs	676
21.6 Miniature Pigs	692
21.7 Rats	696
21.8 Hamster	714
21.9 Mice	714
21.10 Shrew (<i>Suncus murinus</i>)	727
21.11 Rabbit	727
21.12 Sheep Model	728
21.13 Ferrets	730
21.14 Cat	731
21.15 Animal Models of Periodontitis-Associated Systemic Disease	731
21.16 Improving the Reporting of Animal Experiments: The ARRIVE Guidelines	741
21.17 Methodological Quality of Systematic Reviews of Animal Studies	744
21.18 Comparison of Treatment Effects Between Animal Experiments and Clinical Trials	745
References	747
Index	765

Recent clinical research is conducted based on bioethical consideration of human subjects. The Ethical Guidelines for Clinical Studies (EGCS) form the standard for this ‘subject protection’. In current clinical research, consideration of subject rights and life is held more important than the scientific and social value of the research. We describe herein the major revisions and history of ethical considerations leading up to implementation of the revised EGCS on April 1, 2009. The obligations of clinical researchers regarding ethical studies and training and enrolment in insurance for subject compensation have been added to these latest guidelines. The role of ethics review boards, which supervise whether clinical researchers are actively performing subject protection, is also becoming extremely important (Sekine and Shimada 2011).

2.1 Protection of Human Study Participants

2.1.1 Ethical Foundations

Contemporary research standards are robustly encompassed by several guidelines, including the Declaration of Helsinki (DoH) by the World Medical Association (WMA) (Table 2.1). These requirements attempt to protect human rights, autonomy, privacy, and justice and ensure the participants’ benefits/harm, as well as maintaining ethical standards and the integrity of biomedical research on humans (Pitak-Arnnp et al. 2009).

The Nuremberg Code and Declaration of Helsinki are the international ethical standards for human-subjects research. After World War II, the Nuremberg trials were conducted to prosecute Nazi leaders for crimes against humanity. A substantial proportion of the trials involved Nazi physicians who had forced prisoners to undergo appalling, inhumane procedures in the name of clinical research. At the time of the Nuremberg court, there were no laws, regulations, codes, or formal documents that stated ethical standards for human-subjects research, so the trial proceedings resulted in the development of a document, **the Nuremberg Code**, that articulated the basic requirements for conducting research in a manner that respects the fundamental rights of human subjects. The three basic elements of the Nuremberg Code (voluntary and informed consent, a favourable risk-to-benefit analysis, and the right to withdraw without repercussions) became the foundation for subsequent ethics codes and federal research regulations. Thus, every person involved in human-subjects research should read the Nuremberg Code (<http://www.hhs.gov/ohrp/references/nurcode.htm>) (Rice 2008).

Furthermore, three events during the 1960s heralded a change in the ethical oversight of human-subjects research. In the late 1950s, the investigational drug thalidomide was used to treat discomforts associated with pregnancy, including morning sickness and insomnia. At the time it was neither a requirement nor standard practice to inform patients of the investigational nature of pharmaceuticals being tested. In 1962, however,

Table 2.1 International guidelines concerning human subject protections. (Pitak-Arnnop et al. 2009) (Reprinted with permission from Elsevier)

Guidelines	Websites
Nuremberg Code (Directives for Human Experimentation)	http://ohsr.od.nih.gov/guidelines/nuremberg.html
Declaration of Helsinki (DoH) of the World Medical Association (WMA)	http://www.wma.net/e/policy/b3.htm
International Ethical Guidelines for Biomedical Research involving Human Subjects of the Council for International Organisations of Medical Sciences (CIOMS) and the World Health Organisation (WHO)	http://www.cioms.ch/frame_guidelines_nov_2002.htm
Universal Declaration on Bioethics and Human Rights of the United Nations Educational Scientific and Cultural Organisation (UNESCO)	http://unesdoc.unesco.org/images/0014/001461/146180E.pdf
Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine of the Council of Europe	http://conventions.coe.int/treaty/en/treaties/html/164.htm
Uniform Requirements for Manuscripts to Biomedical Journals (updated in 2007) of the International Committee of Medical Journal Editors (ICMJE)	http://www.icmje.org/
Recommendations on Publication Ethics Policies for Medical Journals of the World Association of Medical Editors (WAME)	http://www.wame.org/resources/publication-ethics-policies-for-medical-journals
Guidelines on good publication and the Code of Conduct for Editors of Biomedical Journals of the Committee on Publication Ethics (COPE)	http://www.publicationethics.org.uk/guidelines

it became apparent that thalidomide caused birth deformities. Public outrage about these foetal deformities resulted in an amendment to the Food, Drug, and Cosmetic Act; investigators were required to obtain informed consent from subjects before administering investigational medications. In 1964, the World Medical Association met in Helsinki, Finland, to draft a document that describes the ethical standards of human-subjects research. In addition to the three central principles of the Nuremberg Code, **the Declaration of Helsinki** (<http://www.ftsr.ulaval.ca/ethiques/DOH/1975.pdf>) added two novel elements:

- The interests of the subject should always be placed above the interests of society.
- Every subject should get the best known treatment (Rice 2008).

The Declaration remained untouched until 1975, although on several occasions, there were attempts or suggestions to have it revised in light of the rapid advance of medical technology. Although the general focus and core ideals remained the same, the 1975 revisions were extensive, that is, changing the terminology through-

out, adding 17 new paragraphs, amending several existing points, and restructuring the document. The result of these changes was that the list of ‘basic principles’ was expanded, and the following two categories (the first of which deals with research combined with therapeutic care and the second with research for purely scientific purposes) were crafted to provide guidelines more specific to those particular circumstances. In 1983, all WMA Declarations and Statements were reviewed to ensure consistency and the use of up-to-date terminology. The Declaration of Helsinki was revised again, although most of the revisions were strictly editorial in nature. The only substantive change made was the addition of a clause stating that ‘when a child is to be a subject for research, that minor’s consent must be obtained’. In 1989, the German Medical Association expressed concern about the declaration’s Basic Principle 2 which calls for an experimental protocol to be transmitted to a ‘specially appointed independent committee’. The German Medical Association introduced an amendment to define more clearly the appointment and status of this

committee which was adopted at the assembly in Hong Kong, in 1989. The declaration was again amended at the 48th assembly in South Africa, in 1996. This time, the revision represented the addition of a new sentence in order ‘not to exclude the use of inert placebo in studies where no proven diagnostic or therapeutic method exists’ (WMA 2011). Further changes were done at the 52nd WMA General Assembly, Edinburgh, Scotland, October 2000; 53rd WMA General Assembly, Washington 2002 (Note of Clarification on paragraph 29 added); 55th WMA General Assembly, Tokyo 2004 (Note of Clarification on Paragraph 30 added); and 59th WMA General Assembly, Seoul, October 2008. The 2008 form can be read on the WMA web page at <http://www.wma.net/en/30publications/10policies/b3/17c.pdf>.

The only other document so appended in the case of the ICH Guideline is the well-known Belmont Report, that is, the April 1979 Report on ‘Ethical Principles and Guidelines for the Protection of Human Subjects of Research’, formulated by the (US) National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The new European Union directive on GCP for clinical trials (2001/20/EC), which was published on May 1, 2001, includes a preambular paragraph 2 affirming that ‘The accepted basis for the conduct of clinical trials in humans is founded in the protection of human rights and the dignity of the human being with regard to the application of biology and medicine, as for instance reflected in the 1996 version of the Helsinki Declaration’ (Human and Fluss 2001; Bhutta 2002).

In United States, the National Research Act (Pub. L. 93–348) was signed into law, in 1974, creating the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report, ‘Ethical Principles and Guidelines’, summarizes the commission’s work and includes respect for persons, beneficence, justice, application of the general principles in informed consent, assessment of risk and benefit, and selection of subjects. **Respect for persons** acknowledges that each individual should be treated as an autonomous agent and that those with diminished autonomy

are entitled to protection. **Beneficence** refers to respecting a subject’s decisions and protecting a subject from harm. This includes maximizing benefits and minimizing harm, actual, or potential. **Justice** incorporates fairness of distribution so that no one is denied a benefit unless there is good cause. An investigator’s number one responsibility is to design and implement research based on these three principles. Application of these three briefly described principles is evidenced in the areas of informed consent, risk/benefit assessment, and selection of subjects for research (Belmont Report: Office of Human Subjects Research 2011; Erler and Thompson 2008). Federal law requires the Department of Health and Human Services (DHHS) to issue regulations for the protection of human subjects of research and to implement a program of instruction and guidance in ethical issues associated with research. These regulations are codified at Title 45, Part 46 of the Code of Federal Regulations, Protection of Human Subjects, 45 CFR 46 (Responsibilities of the Office of Human Subjects Research 2011; Erler and Thompson 2008).

In 1997, the International Committee of Medical Journal Editors (ICMJE) declared the ‘Uniform Requirements for Manuscripts to Biomedical Journals’ (updated in 2008). It stated that authors should disclose whether their research procedures follow international and national ethical standards and the DoH. The DoH specifically requires a comprehensive and systematic approach to the scientific, ethical, legal, and social issues of research by an independent, multidisciplinary research ethics committee (REC) or institutional review board (IRB) and freely given informed consent from the subjects. The ICMJE requirements for human protection are equivalent to both components of the DoH. The ICMJE guideline has been adopted by more than 500 journals and is often included as part of the instructions for manuscript preparation. This guideline is also consistent with the recommendations by the World Association of Medical Editors (WAME) and the Committee on Publication Ethics (COPE) (Pitak-Arnnp et al. 2009). Ethical requirements are clearly written in the journals’ guidelines for authors (Table 2.2).

Table 2.2 Instructions for authors of the three main periodontology journals with regard to ethical requirements

Journals	Ethical prerequisites of manuscripts
Journal of Periodontology	Most case reports and case series are a retrospective description of clinical findings in a case or an observed course of events that document a new aspect of patient management during the normal course of clinical treatment. Since there is no hypothesis testing, no systematic data collection beyond that which is part of routine clinical practice, no data analysis, and the work has already been done; case reports and case series do not usually qualify as 'research' requiring approval from ethical boards designed to protect humans involved in clinical research. (US Fed. definition: 'RESEARCH is any systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge'.) Example 1: Series of private practice implant cases in patients who have been taking bisphosphonates. Authors describe the findings in each case, which are collected and reported in a table format. Example 2: Authors collect series of private practice implant cases in patients who have or have not been taking bisphosphonates. The sample size is sufficient for data analysis, and authors analyze and report the incidence of complications. Example 1 does not qualify as 'research', but example 2 does qualify and requires ethical approval. Our policy conforms to the Uniform Requirements, which states 'Patients have a right to privacy that should not be infringed without informed consent. Identifying information should not be published in written descriptions, photographs, and pedigrees, unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication. Informed consent for this purpose requires that the patient be shown the manuscript to be published'.
Journal of Clinical Periodontology	Authorship and Acknowledgements: Authors submitting a paper do so on the understanding that the manuscript has been read and approved by all authors and that all authors agree to the submission of the manuscript to the journal. <i>Journal of Clinical Periodontology</i> adheres to the definition of authorship set up by The International Committee of Medical Journal Editors (ICMJE). According to the ICMJE, authorship criteria should be based on (1) substantial contributions to conception and design of or acquisition of data or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3. It is a requirement that all authors have been accredited as appropriate upon submission of the manuscript. Contributors who do not qualify as authors should be mentioned under Acknowledgements. Please note that it is a requirement to include email addresses for all coauthors at submission. If any of the email addresses supplied are incorrect, the corresponding author will be contacted by the journal administrator. Under acknowledgements, please specify contributors to the article other than the authors accredited. Ethical Approvals: Experimentation involving human subjects will only be published if such research has been conducted in full accordance with ethical principles, including the World Medical Association <i>Declaration of Helsinki</i> (version 2008) and the additional requirements, if any, of the country where the research has been carried out. Manuscripts must be accompanied by a statement that the experiments were undertaken with the understanding and written consent of each subject and according to the above-mentioned principles. A statement regarding the fact that the study has been independently reviewed and approved by an ethical board should also be included. When experimental animals are used, the methods section must clearly indicate that adequate measures were taken to minimize pain or discomfort. Experiments should be carried out in accordance with the guidelines laid down by the National Institute of Health (NIH) in the USA regarding the care and use of animals for experimental procedures or with the European Communities Council Directive of 24, November 1986 (86/609/EEC) and in accordance with local laws and regulations. All studies using human or animal subjects should include an explicit statement in the Material and Methods section identifying the review and ethics committee approval for each study, if applicable. Editors reserve the right to reject papers if there is doubt as to whether appropriate procedures have been used. Conflict of Interest and Sources of Funding: Authors are required to disclose all sources of institutional, private, and corporate financial support for their study. Suppliers of materials (for free or at a discount from current rates) should be named in the source of funding and their location (town, state/county, country) included. Other suppliers will be identified in the text. If no funding has been available other than that of the author's institution, this should be specified upon submission. Authors are also required to disclose any potential conflict of interest. These include financial interests (e.g. patent, ownership, stock ownership, consultancies, and speaker's fee) or provision of study materials by their manufacturer for free or at a discount from current rates. Author's conflict of interest (or information specifying the absence of conflicts of interest) and the sources of funding for the research will be published under a separate heading entitled 'Conflict of Interest and Sources of Funding Statement'. See Editor-in-Chief Maurizio Tonetti's editorial on Conflict of Interest and Sources of Funding and www.icmje.org/#conflicts for generally accepted definitions.

Journal of Periodontal Research	The <i>Journal of Periodontal Research</i> adheres to the below ethical guidelines for publication and research. Authorship and Acknowledgements: Authors submitting a paper do so on the understanding that the manuscript has been read and approved by all authors and that all authors agree to the submission of the manuscript to the journal. The <i>Journal of Periodontal Research</i> adheres to the definition of authorship set up by The International Committee of Medical Journal Editors (ICMJE). According to the ICMJE, authorship criteria should be based on (1) substantial contributions to conception and design of or acquisition of data or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3. It is a requirement that all authors have been accredited as appropriate upon submission of the manuscript. Contributors who do not qualify as authors should be mentioned under Acknowledgements. Acknowledgements: Under acknowledgements, please specify contributors to the article other than the authors accredited. Acknowledge only persons who have made substantive contributions to the study. Authors are responsible for obtaining written permission from everyone acknowledged by name because readers may infer their endorsement of the data and conclusions. Please also include specifications of the source of funding for the study and any potential conflict of interests if appropriate. Suppliers of materials should be named and their location (town, state/county, country) included. Ethical Approvals: All studies using human or animal subjects should include an explicit statement in the Material and Methods section identifying the review and ethics committee approval for each study, if applicable. Editors reserve the right to reject papers if there is doubt as to whether appropriate procedures have been used. Conflict of Interest and Source of Funding: Please disclose information concerning sources of institutional, private, and corporate financial support for the work within the manuscript be fully acknowledged and any potential conflicts of interest under Acknowledgements
Clinical Advances in Periodontics	Requirement for Ethics Board Approval: Most case reports and case series are a retrospective description of clinical findings in a case or an observed course of events that documents a new aspect of patient management during the normal course of clinical treatment. If the manuscript includes the data analysis results of testing a specific hypothesis or the treatment or collection of data beyond that which is part of routine clinical management of the case, it qualifies as “research” and must have an institutional review board approval and written informed consent from each patient. The journal adheres to the principles set forth in the Helsinki Declaration of 1975, as revised in 2000, and requires that all reported research involving human subjects be conducted in accordance with such principles. For more information on ethics board approval, please see the “Requirement for Ethics Board Approval” section on page 3 of the <i>Journal of Periodontology</i> Instructions to Authors. Acknowledgments: Acknowledgments should include any contributors who do not meet the standards for authorship, must include the source of any funding for the study, and must define the commercial relationships of each author. Please include affiliations for individuals and locations for funding agencies listed in the Acknowledgments. Authorship: Individuals identified as authors must meet the following criteria established by the International Committee of Medical Journal Editors (http://www.icmje.org): 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; and 3) final approval of the version to be published. Conflict of Interest and Financial Disclosure Forms: A conflict of interest and financial disclosure form must be submitted by each author. A link to this form will be e-mailed to each author after manuscript submission. More information on conflicts of interest can be found under the “Conflicts of Interest” section on pages 5 and 6 of the <i>Journal of Periodontology</i> Instructions to Authors.

2.1.2 Who Judges Whether Research Is 'Ethical'?

In a formal manner, hospital and university ethics committees are appointed specifically for this purpose. However, those who allow access to their settings (unit managers) and consumers/the media also comment on the ethical conduct of research (Endacott 2004). In the United States, Institutional Review Boards (IRBs) were established by federal legislation in response to growing concerns over ethics in human-subjects research, especially with vulnerable subjects. The main role of IRBs is to protect the rights and welfare of human subjects. The complexity of providing such protection has rendered the makeup and modern-day functioning of IRBs complex. The ultimate goal of the IRB, however, parallels that of the researcher; both are charged with ensuring that human-subjects research is conducted ethically, with sound scientific rationale, to maximize benefits and minimize risks (Rice 2008). However, the 2008 form of Declaration of Helsinki states that 'Physicians should consider the ethical, legal and regulatory norms and standards for research involving human subjects in their own countries as well as applicable international norms and standards. No national or international ethical, legal or regulatory requirement should reduce or eliminate any of the protections for research subjects set forth in this Declaration'.

The key concerns of ethics committees can be grouped under three headings (Endacott 2004):

- *Is this 'good science'?* This issue is often debated by recipients of critical reports from ethics committees, but a key ethical concern is that participants shouldn't waste their time on poorly designed research. The competence of the researchers to undertake the study is also assessed by the ethics committee.
- *What are the benefits, costs, and risks for patients/participants?* It is generally accepted by participants that their input may not yield benefit for them personally. However, researchers should bear in mind the balance of cost and benefit. It may be appropriate to deliver a seminar to present and discuss findings in the site

where data collection was undertaken. It is common practice to identify what processes are in place for dealing with any potential harm, physical or psychological, that may result to the patient. This usually requires identifying the trigger points for reporting or intervening if bad practice is revealed through the research data collection (e.g. if patient safety is compromised). A useful strategy is to refer any matters of concern through the hospital clinical risk management processes. Drug trials have detailed procedures for reporting and managing any harm to participants.

- *What are the benefits, costs, and risks for this hospital/department?* Costs to be considered by ethics committees include the time to be taken for staff in providing data or in attending briefing sessions regarding the research. In the UK, there are processes in place at NHS trust level to identify all such 'hidden' costs. Ethics committees also seek to protect both staff and patients from being 'over-researched'.

Ethical review is the responsibility of every institute that participates in a project. The increase in multicentred diagnostic and therapeutic trials in medicine and dentistry over the past two decades has put unanticipated pressure on the REB system. A single project can require ethical review by dozens of independent institutes, and their various REBs may come to different conclusions. The effect on time-sensitive research programs of unanticipated delays in obtaining multiple ethical approvals has led some clinical research investigators to question whether imposing unnecessary delays on ethical research is in itself unethical (Nicholl 2000; Ellen and Singleton 2008).

When experimentation in practice is a part of a formal community research network, who has responsibility for conducting the ethical review? We assume that the practitioner would be appointed to a sponsoring hospital, university, or research institute team, but we have been unable to locate a document that recommends a specific formal process. In many cases, practitioners report the analyzed results of treatment outcomes from their own practice in peer-reviewed journals. For example, some analyses of long-term maintenance care in private settings are

considered ‘classics’ in the periodontology literature. This type of research should certainly comply with the ethical review procedures outlined in government policies. Yet, it is unclear how arm’s length ethical review of private practice-based research is currently conducted. Perhaps this is an issue that should be debated by professional associations and licensing bodies (Ellen and Singleton 2008).

Human-subject protection remains a gray area in some circumstances, such as the use of human tissue or specimens, audit or quality assurance, educational research (such as in students or trainees), and a case report or series. There are considerable differences between countries (Table 2.3). The US federal courts have denied individuals the right to own biological material, including pathological specimens after surgery, contributed for research. This may be a reason why many reports from the USA rarely mentioned ethical protection if patient data or records and/or pathological or diagnostic specimens were analyzed. There are also variations in the ethical review procedures of REC, within and between countries, and these may influence research practice, especially in conducting a multicenter study (Pitak-Arnnp et al. 2009).

2.1.3 Informed Consent

Consent is understood differently by various disciplines and professions and also in various theoretical models. Consent is a strong concept in being so versatile and durable, but it is vulnerable to conflicting interpretations and rejection as a worthless ideal. Real and critical consent reminds practitioners and researchers about standards that protect them and their patients; these are at times too high to achieve, but they endure as standards to aim for. Functionalist consent reflects common medical practice, if not medical values. Social construction shows how consent is a process—perceived, experienced, and shaped through interactions between individuals and their social contexts. Postmodernism provides revealing descriptions of current contradictions and confusions in consent, which is usually assumed to

Table 2.3 Differences among three countries in relation to ethical approval by a research ethics committee (REC) or institutional review board (IRB) (Pitak-Arnnp et al. 2009) (Reprinted with permission from Elsevier)

UK^a
The Central Office for Research Ethics Committees (COREC) requirements for ethical review
1. patients or their relatives or carers and users of the National Health Service (NHS)
2. access to data, organs, specimens or other bodily materials of past and present NHS patients
3. foetal materials and <i>in vitro</i> fertilisation of NHS patients
4. the recently dead in NHS premises
5. the use of NHS premises or facilities
6. NHS staff recruited as research participants by virtue of their professional role
USA^b
Federal regulations’ exceptions (‘not’ require ethical review)
1. research involving normal educational practices or tests
2. survey or interview procedures
3. an investigation of existing data, documents or records, pathological or diagnostic specimens
France^{c,d}
Research types which require ethical review
1. retrospective analyses of specimens
2. all prospective studies (except prospective observational studies without an invasive procedure on the human body)
Federal regulations’ exceptions
1. an investigation on patient data or records
2. a survey of healthcare providers

Sources:

^a<http://www.webarchive.org.uk/pan/13880/20070125/www.corec.org.uk/applicants/help/faqs.htm#approval>

^b<http://www.hhs.gov/ohrp/documents/OHRPRegulations.pdf>

^c<http://infodoc.inserm.fr/inserm/ethique.nsf/937238520af658aec125704b002bded2/d58ce05b3098a503c12570fa004e750f?OpenDocument>

^dhttp://www.drcc.aphp.fr/encadrement_juriregl/introduction.php

epitomise rational and moral certainties. Consent is too complex to be understood fully in any one theoretical model (Alderson and Goodey 1998).

Informed consent is important. Fundamentally, the reason why it is important is not that it is a regulatory requirement but that it serves a function: to allow subjects to make an informed and

voluntary choice to participate—or refuse to participate—in a project where they will be asked to take risks for the benefit of others. In both research and clinical care, informed consent represents a permission to intervene on a person's private sphere, especially when a physical intervention is involved (Cahana and Hurst 2008).

Although informed consent is considered crucial both for research participation and for clinical care, these two types of situations are not identical. In research, it allows subjects to make an informed and voluntary choice to participate—or refuse to participate—in a project where they will be asked to take risks for the benefit of others (Cahana and Hurst 2008). Consequently, the usual informed, specific consent that a doctor receives from a patient may not be applicable to ventures that combine research and health-care goals. Iceland, for example, provides assumed or presumed consent with a provision for people to opt out. Estonia asks their citizens for open consent when they provide blood samples and health-care information to the database (Kegley 2004). Commonly, in both research and clinical care, informed consent represents a permission to intervene on a person's private sphere (Cahana and Hurst 2008).

2.1.3.1 The Elements of Adequate Informed Consent

The elements of adequate informed consent include disclosure, understanding, decision-making capacity, and voluntariness. Application of each of these elements requires a degree of judgment by the clinician/researcher; however, each poses distinct difficulties. Promoting informed and voluntary consent benefits from knowledge of these obstacles and of methods effective in overcoming them. Consequently, the degree to which informed consent fulfills each of these elements and how each can be affected by interventions designed to improve them has become a specific area of research (Cahana and Hurst 2008).

Disclosure

Current practice in obtaining informed consent seems to have been shaped by emphasis on the legal duty of disclosure, particularly in the

research setting. Consent is seen as an action, concluded by signing a form. However, informed consent needs not only disclosure and a signature but also promotion of participants' understanding of the research project and the voluntary nature of their decision to participate (Jefford and Moore 2008; Brody et al. 2005).

The written informed-consent document (i.e. consent form) is an important part of the requirement to disclose and advise participants of the details of a proposed trial. Although the form has been said to give 'legal and symbolic documentation of an agreement to participate' (Jefford and Moore 2008; Silverman et al. 2005), the length and complexity of informed-consent documents hinder participant understanding (Jefford and Moore 2008; Jefford et al. 2005). Viewing the consent form mainly as a legal document tends to hinder attempts to create reader-friendly documents: 'many sponsors and institutions appear to view them primarily as a legal instrument to protect them against litigation' (Jefford and Moore 2008).

Understanding

Appropriate information given to a competent individual will promote understanding and, in this regard, sensible decision-making without coercion. However, the process of understanding involves the interaction of psychological and intellectual characteristics of an individual and depends on the educational status, the level of general knowledge, and personal attitudes, which are affected by the morals and customs of the society. The communication of medical information to patients is even more demanding because of the need to explain scientific issues with plain language. Physicians should also communicate medical information in a caring and compassionate way. If these requirements are met, the clinician–patient relationship will be founded on trust and alliance. A recent review evaluated the degree of patients' understanding of several aspects of the informed consent process for surgery and clinical research. Regarding surgery, adequate overall understanding of the information provided and of the risks associated with surgery was shown in 29% and 36% studies providing

relevant data, respectively. Regarding clinical research, adequate understanding of the aim of the study, the process of randomization, voluntarism, withdrawal, and the risks and the benefits of treatment was shown in 54%, 50%, 47%, 44%, 50%, and 57% of studies providing relevant data, respectively. Satisfaction by the amount of the given information was shown in 58% studies involving surgery and 80% studies involving clinical research (Falagas et al. 2009).

Decision-Making Capacity

Informed consent requires a voluntary and informed decision by a competent person. The assessment of a person's capacity for consent can be conceptualized as a two-stage process. First, the abilities relevant to decision-making are measured. However, an instrument score reflects only one factor in the capacity evaluation. The second step is the clinical judgment that incorporates the subject's performance factors along with important contextual factors—the main one being the risk/benefit calculus—to yield a dichotomous decision. The first task requires an appraisal of the cognitive abilities relevant to capacity and is focused on the patient/subject; the second judges this information in a specific context to yield a categorical decision regarding 'capacity'. For example, Alzheimer's disease is identified by the National Bioethics Advisory Commission as one of the 'mental disorders that may affect decision-making capacity'. The current practice involving most potential subjects with probable Alzheimer's disease involves a 'double consent' procedure for clinical trials, that is, consent or assent from the Alzheimer's disease subject along with a surrogate consent (Kim et al. 2001).

Respect for vulnerable people and voluntary, informed consent requires special consideration in oral health research. In Canada, dental diseases are prevalent among Aboriginal populations, the elderly, the poor, and those without dental insurance. Serious ethical concerns revolve around informed consent for children and for medically compromised adults who are unable to provide their own consent (Ellen and Singleton 2008; Verástegui 2006). Clinical trials in developing countries present additional ethical dilemmas (Lo 2006).

Voluntariness

Another premise of consent is that of volunteerism, which implies that after being fully informed, subjects choose to participate, without coercion (Erler and Thompson 2008; Macklin 1999).

Determining how much detail should be provided and balancing this with the potential participants' need for information and capabilities to understand, it is a major challenge for researchers and institutional review boards (Rivera et al. 2007). To facilitate these efforts, the US Code of Federal Regulations (CFR) (Office for Protection from Research Risks 1996) requires that the following eight basic elements be presented to participants to ensure informed consent: (1) research description, (2) risks, (3) benefits, (4) alternatives to study participation, (5) confidentiality, (6) compensation, (7) contacts, and (8) voluntary participation. The CFR also provides a brief elucidation of what type of information should be provided under each element. This limited clarification of the CFR creates the potential for different interpretations; more detailed specifications should be considered (Rivera et al. 2007).

2.1.3.2 Improving the Informed Consent

Administration of the consent form involves a verbal explanation of the study, including purpose, number of participants, study procedure, risks and benefits of participation, confidentiality, cost or compensations, the voluntary nature of participation, which addresses the subject's right to refuse participation or to withdraw without penalty, and contact information for the researcher. The consent form should (1) be written at a reading level that potential subjects will be able to understand (typically sixth to eighth grade), (2) provide simple definitions, avoiding use of jargon, and (3) use subject headings and an easy-to-read font. The researcher is responsible for assuring that the subject is competent, understands the study content presented in the consent form, and is not under any undue stress that might influence participation. A subject's signature is obtained to document verification of the explanation (Erler and Thompson 2008).

Comprehension requires that the patient be able to understand the information presented and have the time and opportunity to read, evaluate, and consider the information presented. Readability of the consent form, while contributing to the comprehension, is not the same as comprehension. An easy-to-read text might be difficult to comprehend if poorly written. However, readability does affect the willingness to read the text and hence could improve comprehension. The FDA has stated the responsibility of the Institutional Review Board (IRB) as ‘The IRB should ensure that the informed consent document properly translates complex scientific concepts into simple concepts that the typical subject can read and comprehend’ (Pandiya 2010).

Plain language aims to simplify explanations such that the meaning is clear to a readership with varying reading experience and abilities. A clear and simple message is likely to be understood by more people than a message with complex wording. Plain language is important during the consent discussion. The discussion allows for checking of understanding—an important advantage over a form; however, simply it is written. **Table 2.4** is a checklist for clear communication (Jefford and Moore 2008).

A simple mechanism of reducing the reading difficulty level is by replacing, wherever possible, technical terms with common terms. Using tools like outlining, bullet points, a large typeface, and diagrams can help the reader to follow complex concepts. Using active verbs rather than the passive voice, short sentences, and frequent paragraphing can make text simpler. Several computer software packages allow rapid analysis of the readability score. Most ethics committees recommend a readability score of grade 8. However, often the actual readability may differ from the prescribed standard. Besides, the readability may not reflect appropriately on the comprehensibility of the consent form (Pandiya 2010; Michael et al. 2003).

Multimedia and enhanced consent form interventions do not consistently improve research participants’ understanding. Person-to-person interactions, especially the extended

Table 2.4 Proposed checklist for clear communication (Jefford and Moore 2008) (Reprinted with permission from Elsevier)

• Use familiar words and ideas
• Use short words and sentences where possible
• Avoid misleading descriptions
• Use readability checkers or formulae to estimate reading grade level
• Discuss standard treatments and trial treatments
• Encourage support from others
• Give information to take away
• Check understanding

discussion interventions, may be more effective in improving understanding (Flory and Emanuel 2004).

Enhancing print through the use of descriptive headings, white space, simplified language, and illustrations was the best way to convey information to low-income, minority patients. Arguably, if non-print media are used to convey information to poor readers, it might be well to reinforce the message with written materials. An advantage of written material is that it can be taken home for later review. It was sobering to learn that only about half the information conveyed was demonstrated through immediate recall, and then, only if prompted. This reinforces the need for investigators to make sure that study participants truly do understand the implications of taking part in research (Campbell et al. 2004).

The length of the consent form is another issue that needs to be examined. Many consent forms are 15–20 pages long. The length by itself might act as a deterrent for the form to be read. Besides, to simply read a document that is 20 pages long would mean a time spent of approximately 60 min. In a clinical setting, this can be a considerable amount of time, and the total time spent in the consenting process might be reduced, with the patient’s queries not getting addressed adequately. Besides the amount of time spent, too much of information may hinder the understanding of information that is relevant to the patient. It is no surprise that in studies done on the informed consent form, it was found that patients prefer simpler and easier to read informed consent forms that can provide them the necessary

information to make a decision regarding participation in the trial (Pandiya 2010; Dresden and Levitt 2001; Brody 2001 Paasche-Orlow et al. 2003).

2.1.3.3 Side Effects

Giving more information to patients and research subjects is sometimes perceived as a potential hazard to them, and concerns are sometimes voiced as to whether it might increase anxiety or decrease consent to clinical interventions or accrual in research. In examining these issues, it is crucial to distinguish two elements. First, the presence—or absence—of side effects, such as patient anxiety, does not automatically lead to a conclusion about whether clinicians should disclose relevant information. Patients may prefer to receive information even if it makes them anxious and have the right to be offered information relevant to personal choices even when they also have the right to refuse such information. Second, the very purpose of requiring informed consent is to allow patients and subjects to refuse their consent, should they so wish. Side effects, such as lower acceptance or accrual rates, thus, cannot represent a reason to give less information: they simply reflect the fact that patients and subjects are exercising this right. However, it is still useful to know whether we should expect such effects or not. A clinical investigator, for example, would benefit from knowing whether better informed consent requires him to plan a longer enrollment period for his research (Cahana and Hurst 2008).

2.2 Responsibilities of Investigators

2.2.1 Conflicts of Interest

In law, the term conflict of interest is used primarily in connection with fiduciaries. A fiduciary holds some form of power that is to be used for the benefit of another, based on specialized knowledge or expertise. The fiduciary relationship involves dependence, reliance, and trust and is held to the highest legal standard of conduct (Morin et al. 2002).

Conflicts of interest are inherent to the majority of relationships among individuals and of these with companies and institutions, and certainly, research involving human beings is no exception. In relation to clinical research, conflicts of interest occur at different levels and usually permeate among them: in the pharmaceutical industry in their decisions to invest to develop new products, especially vaccines and drugs, and also in relation to marketing of these products; among the investigators, the conflicts may be related to the financial gains to participate in pharma-sponsored trials, or to the expected academic career boost attained with the publication of the results of the trials and also to personal interests such as the financial support for trips to international conferences. Often the participation of host country investigators is restricted to performing phase III or IV protocols developed abroad, many times with low scientific relevance, and even lower relevance to public health; universities or research institutes themselves also have conflicts of interest, as the sponsored projects may help increase their budgets, both directly (taxes) and indirectly (e.g. improvement of physical infrastructure of laboratories or outpatient clinics); for the trial volunteers in developing countries, participation in clinical trials is many times seen as, and can really be, an unique opportunity of receiving better health care, better treatment by the health professionals, and easier access to costly lab exams and also to receiving certain medications which would otherwise be difficult to have access to (Greco and Diniz 2008).

Investigators can minimize conflicting interests in their clinical trials. For example, blinding investigators and participants to which intervention the participant is receiving and using an independent data safety monitoring board can prevent bias in assessing outcomes and interpreting interim data. For industry-sponsored research, investigators should have control over the primary data and statistical analysis and the freedom to publish findings whether or not the drug or device is found to be effective. Physician investigators should also discuss their research plans with colleagues. Formal and informal peer review provides an excellent mechanism for evaluating the risks and benefits of research and identifying areas of concern (Wolf and Lo 2000).

For people to make informed decisions about participating in a clinical trial, physicians need to disclose pertinent conflicts of interests. In a landmark court case, the California Supreme Court declared that physicians need to ‘disclose personal interests unrelated to the patient’s health, whether research or economic, that may affect the physician’s professional judgment’. Some individuals more closely scrutinize a study in which the investigators have a direct financial stake. Furthermore, disclosures of conflicts of interests are salutary because they deter researchers from entering into questionable financial arrangements that would be difficult to justify to the public or to their peers (Wolf and Lo 2000).

2.2.2 Investigator Integrity

Scientific misconduct and conflict of interest differ in a very basic way. Scientific misconduct clearly requires some action on the part of the investigator which falls under the recognized categories of fabrication or falsification of data and plagiarism. In contrast, conflict of interest can be considerably more subtle since it arises as a result of a set of circumstances which may provide the potential for unethical conduct on the part of the investigator but which in and of itself should not be considered unethical. Examples of such circumstances include situations in which the investigator of a product also holds an equity position in the company producing the product (and thereby may have an incentive to emphasize positive data while minimizing less favourable data), situations in which a lecturer or writer on a specific topic may receive considerable research and/or consultant funding from a company marketing a product within the topic category of the lectures or papers, or other relationships between a researcher and corporation which may result in financial gain to the researcher. In all these instances, ethical questions arise around the issue of whether these circumstances will actually influence the judgment of the investigator in such a manner as to place the desire for personal gain ahead of the stan-

dards of objectivity and intellectual honesty that cause the investigator to place personal interests above those of professional and societal interests (Barnett 1995).

To sum up, therefore, while scientific misconduct is characterized by actual behaviours which deviate from the norms of ethical scientific conduct, conflict of interest is defined by a set of circumstances which provides the potential for unethical conduct resulting in a compromise of the integrity of the research process but is not in itself unethical conduct per se (Barnett 1995).

Research investigators have an ethical obligation to conduct their research honestly through the judicious use of grant and contract funds for the purposes intended; accuracy in fully disclosing all research strategies, methods, results, and analyses; and generous and accurate citation of other investigators’ preceding or competing work. For whatever reason or pressure—career advancement, financial gain, personal fame, or competitive spirit—some investigators are drawn into the dishonest practices of falsifying research findings, withholding data, making false claims or plagiarism (Ellen and Singleton 2008).

When reviewing ethical issues that may affect dental research, the former editor of the *Journal of the American Dental Association* identified a seven-tier defense against scientific misconduct: (1) Personal integrity, as most people, given a choice, want to take pride in their work; this includes holding up their heads as honest, unbiased individuals, while poor-quality work, especially with taint of bias, will damage reputation among peers. (2) Right to publish. When practitioners, universities, or other third parties perform research, it is common for the investigator to be contractually guaranteed the right to publish the results. This means that negative findings can’t be unilaterally suppressed. (3) Regulations. In the United States and other developed countries, almost anyone conducting research (including privately sponsored health-related research) must comply with a panoply of government regulations. Many of these involve financial controls, patient safety, and animal welfare. (4) The US Food and Drug

Administration. This layer only applies in the United States (though comparable oversight exists elsewhere) and only to research on drugs and devices that come under FDA jurisdiction. (5) Peer review. Reviewers help the author convey worthwhile research clearly; just as importantly, they sometimes let the editor know when the research is unsuitable for publication. (6) Intelligent readers are trained to determine plausibility and who have an impressive memory of preceding publications. (7) The scientific method. Publication of methods and results invites others to repeat the experiment. What can't be confirmed is consigned to the dustbin of history (Jeffcoat 2002).

2.2.3 Authorship

Jane Austen wrote every word of her novels and was indisputably the author of her books. But conducting scientific research often depends on large teams of people with different skills. Deciding who should be listed as an author is not simple, and too often the decision is made on the basis of power. The powerful are included, even when they have done nothing, and the weak are excluded, even those who have done most of the work. This unethical behaviour can become a major problem if the study proves to be fraudulent, as has happened many times. For authorship implies not just credit, which authors love, but also accountability, about which they are much less enthusiastic (Smith 2006).

Inappropriate authorship may involve **honorary authors**, individuals who are named as authors but who have not met authorship criteria and have not contributed substantially to be able to take public responsibility for the work, and **ghost authors**, individuals who have made substantial contributions to the work reported in an article but who are not named as authors (Wislar et al. 2011). Previous research has documented prevalences of honorary and ghost authors of 19% and 11%, respectively, in articles published in *Biomedical Journals* in 1996 (Flanagin et al. 1998) and of 39% and 9%, respectively, in review articles published by the *Cochrane Library* in

1999 (Mowatt et al. 2002). A recent survey sampled of corresponding authors of 896 research articles, review articles, and editorial/opinion articles published in six general medical journals with high impact factors in 2008: *Annals of Internal Medicine*, *JAMA*, *Lancet*, *Nature Medicine*, *New England Journal of Medicine*, and *PLoS Medicine*. The prevalence of honorary authorship was 25.0% in original research reports, 15.0% in reviews, and 11.2% in editorials, whereas the prevalence of ghost authorship was 11.9% in research articles, 6.0% in reviews, and 5.3% in editorials. However, an overall decrease of honorary and ghost authorship in 21% of articles published in major medical journals in 2008 suggests that increased efforts by scientific journals, individual authors, and academic institutions are essential to promote responsibility, accountability, and transparency in authorship and to maintain integrity in scientific publication (Wislar et al. 2011).

The *AMA Manual of Style* is the bible of style and presentation for *Journal of American Medical Association (JAMA)* and its sister publications, the *Archives* specialty journals. The manual is used not only by *JAMA* editorial and production staff but also by research authors, medical writers, editors, proofreaders, and many others involved in the creation and editing of scientific publications. The first section of Chap. 5 provides a thorough update on article authorship. The definition of 'author' has not changed, but the reporting and publication of author contributions to an article is a new step toward ensuring appropriate authorship (Iverson et al. 2007). In *JAMA*, the author-supplied contribution lists are published in the 'Acknowledgment' section for articles reporting original research, including research letters. In addition, the International Committee of Medical Journal Editors recommends that authors of studies funded by an entity with a potential proprietary or financial interest sign a statement regarding their access to the data (Table 2.5) (Christiansen 2008). However, a recent survey demonstrates that detailed information on authorship/contributorship is not publicly available for high-ranked dental journals (Table 2.6) (Faggion 2011).

Table 2.5 ICMJE guidelines for authorship/contributorship (Faggion 2011) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.

Authorship
• Authorship credit should be based on (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3. • When a large, multicentre group has conducted the work, the group should identify the individuals who accept direct responsibility for the manuscript. These individuals should fully meet the criteria for authorship/contributorship defined above, and editors will ask these individuals to complete journal-specific author and conflict-of-interest disclosure forms. When submitting a manuscript authored by a group, the corresponding author should clearly indicate the preferred citation and identify all individual authors as well as the group name. Journals generally list other members of the group in the Acknowledgments. The NLM indexes the group name and the names of individuals the group has identified as being directly responsible for the manuscript; it also lists the names of collaborators if they are listed in Acknowledgments. • Acquisition of funding, collection of data, or general supervision of the research group alone does not constitute authorship. • All persons designated as authors should qualify for authorship, and all those who qualify should be listed. • Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content.
Contributorship
• All contributors who do not meet the criteria for authorship should be listed in an acknowledgments section. Examples of those who might be acknowledged include a person who provided purely technical help, writing assistance, or a department chairperson who provided only general support. Editors should ask corresponding authors to declare whether they had assistance with study design, data collection, data analysis, or manuscript preparation. If such assistance was available, the authors should disclose the identity of the individuals who provided this assistance and the entity that supported it in the published article. Financial and material support should also be acknowledged. • Groups of persons who have contributed materially to the paper but whose contributions do not justify authorship may be listed under such headings as ‘clinical investigators’ or ‘participating investigators’, and their function or contribution should be described—for example, ‘served as scientific advisors’, ‘critically reviewed the study proposal’, ‘collected data’, or ‘provided and cared for study patients’. Because readers may infer their endorsement of the data and conclusions, these persons must give written permission to be acknowledged.

2.3 Ethical Issues Specific to Certain Types of Research

2.3.1 Randomized Clinical Trials (RCTs)

Although randomized controlled trials are the most rigorous design for evaluating interventions (see Chap. 4), they present special ethical concerns because the intervention is determined by chance. The ethical justification for assigning treatment by randomization in that the arms of the protocol are in equipoise (Freedman 1987; Mann 2006; Robinson et al. 2004; Robinson et al. 2005). That is, current evidence does not prove that either arm is superior. If physicians believe strongly that one arm of the trial is superior and can provide the intervention in that arm outside

the study, they cannot in good faith recommend that their patients enter the trial. Also, the participant might not consider the arms equivalent, for example, when the tradeoffs between benefit and adverse effects differ markedly in a comparison between medical and surgical approaches to a disease (Lo 2006).

How can we judge whether or not potential trial participants have an adequate understanding of what random allocation means? One possible criterion is that participants must demonstrate explicit understanding by giving a verbal definition. Another possible criterion is that participants must demonstrate a working understanding by identifying examples of random and non-random allocation methods. If the participants’ overall pattern of responses matches the pattern given by experts whose understanding is

Table 2.6 Topics assessed directly in the author instructions section of dental journals (Faggion 2011) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

Dental journal	Direct link to ICMJE guidelines? ^a	Merit authorship?	Do not merit authorship but merit an acknowledgement?	Do authors need to describe contributions?	Limited number of authors?	Do authors need to take responsibilities?	Score
Journal of Clinical Periodontology	N	Y	Y ^b	N	N	N	2
Journal of Dental Research	Y	N	Y	N	Y	N	3
Oral Oncology	N	N	N	Y	N	N	1
Periodontology 2000	N	Y ^b	Y ^b	N	N	N	2
Journal of Endodontics	Y ^c	N	N	N	N	N	1
Clinical Oral Implants Research	N	Y	Y	N	Y	N	3
Dental Materials	N	N	N	N	N	N	0
Caries Research	N	N	Y	N	N	N	1
Clinical Implant Dentistry and Related Research	N	N	N	N	N	N	0
Community Dentistry and Oral Epidemiology	N	Y	Y ^b	N	N	N	2
Oral Microbiology and Immunology (Molecular Oral Microbiology)	N	Y ^b	Y ^b	N	N	N	2
Clinical Oral Investigations	N	N	N	N	N	N	0
International Endodontic Journal	N	Y	Y ^b	N	N	N	2
Journal of Periodontology	N	Y	Y	N	N	N	2
Journal of Oral Pathology and Medicine	N	Y	Y ^b	N	N	N	2
Journal of Dentistry	N	Y ^{b,d}	N	N	N	N	1
The International Journal of Oral and Maxillofacial Implants	N	Y ^{e,d}	N	N	N	Y	2
Journal of Periodontal Research	N	Y	Y ^b	N	N	N	2
European Journal of Oral Sciences	N	Y	Y	N	N	N	2
Oral Diseases	N	Y	Y	Y ^g	N	N	3
The Journal of the American Dental Association	N	Y ^b	N	Y ^g	N	N	2
The International Journal of Periodontics and Restorative Dentistry	N	Y ^{e,d}	N	N	N	Y	2
Operative Dentistry	N	N	N	N	N	N	0
Archives of Oral Biology	N	Y	Y	N	N	N	2
The Journal of Adhesive Dentistry	N	N	Y ^b	N	N	N	1
Orthodontics and Craniofacial Research	N	Y	Y ^b	N	N	N	2
Journal of Oral and Maxillofacial Surgery	N	N	Y ^b	N	N	N	1

(continued)

Table 2.6 (continued)

Dental journal	Direct link to ICMJE guidelines? ^a	Merit authorship?	Do not merit authorship but merit an acknowledgement?	Do authors need to describe contributions?	Limited number of authors?	Do authors need to take responsibilities?	Score
Implant Dentistry	N	N	N	N	N	N	0
Oral Surgery Oral Medicine Oral Pathology Oral Radiology and Endodontology	N	Y	Y ^b	N	N	N	2
Journal of Oral Rehabilitation	N	Y	Y ^b	N	N	N	2
International Journal of Oral and Maxillofacial Surgery	N	Y	Y	N	Y ^c	Y	4
Acta Odontologica Scandinavica	N	N	N	Y ^g	N	N	1
American Journal of Orthodontics and Dentofacial Orthopedics	N	N	N	N	N	Y	1
British Journal of Oral and Maxillofacial Surgery	N	Y	Y	N	Y ^c	N	3
Dental Traumatology	N	Y	Y ^b	N	N	N	2
American Journal of Dentistry	N	N	N	N	N	N	0
Journal of Orofacial Pain	N	Y ^{e,d}	Y ^b	N	N	Y	3
Journal of Cranio-Maxillo-Facial Surgery	N	Y	Y	N	N	N	2
Dento Maxillo Facial Radiology	N	N	N	N	N	Y	1
The International Journal of Prosthodontics ⁱ	N	N	Y ^b	N	N	N	1
Australian Dental Journal	N	Y ^b	Y ^b	N	N	Y	3
The Journal of Prosthetic Dentistry	N	N	Y ^b	Y ^g	Y	N	3
International Journal of Paediatric Dentistry	N	Y	Y ^b	N	N	N	2
British Dental Journal	N	N	N	N	N	N	0
Journal of Dental Education	N	N	N	N	N	N	0
European Journal of Dental Education	N	Y	Y ^b	N	N	N	2
Gerodontology	N	Y	Y ^b	N	N	N	2
European Journal of Orthodontics	N	Y	N	N	N	N	1
Community Dental Health	N	N	N	N	N	N	0
Journal of Public Health Dentistry	N	Y	Y ^b	N	N	N	2
Journal of the Canadian Dental Association	N	N	Y ^b	N	N	N	1
Angle Orthodontist	N	N	N	N	N	N	0
Dental Materials Journal ^a							
Swedish Dental Journal	N	N	N	N	N	N	0
Journal of Orofacial Orthopedics	N	N	N	N	N	N	0

Cleft Palate-Craniofacial Journal	N	Y	N	N	N	Y	2
Journal of Esthetic and Restorative Dentistry	N	N	Y	N	N	N	1
Odontology	N	N	N	N	N	N	0
Quintessence International	N	Y ^{d,e}	N	N	N	Y	2
International Dental Journal	N	Y	Y	N	Y	N	3
Journal of Cranio-Mandibular Practice	N	N	N	N	N	N	0
Journal of Applied Oral Science	N	N	N	N	N	Y	1
Revue de Stomatologie et de Chirurgie Maxillo-Faciale ^h							
Journal of Dental Sciences	N	N	Y	N	N	N	1

^aThe Author Instructions section of the journal contains a direct link to the ICMJE authorship and contributorship sections, which provide full instructions (http://www.icmje.org/ethical_1author.html)

^bThe journal indicates that only authors who contributed to the papers should sign as authors. Other individuals should merit an acknowledgment. Nevertheless, the journal does not describe examples of authorship/contributorship

^cA link is provided to the main homepage of the ICMJE, but not to the authorship section

^dThe journal requires that the authors sign a form, but there are no guidelines for authorship/contributorship

^eThe journal reports that 'only individuals who made a significant scientific contribution to the study' should be reported as authors

^fLimited only for some type of papers

^gEditors might ask authors to describe the role of every author in the paper, but this information is not made public

^hInformation not available

ⁱThe Author Instructions section reports that authors need to send a signed form from the publisher's homepage, but there is no address or link to this form

not in doubt, then we can infer that the participants have an adequate working understanding of random allocation. This kind of distinction between explicit and working (or implicit) knowledge is widely used in cognitive psychology—for example, in connection with people’s ability to use syntactical rules without being able to state what those rules are. It has been argued that working understanding of random methods of allocation, rather than explicit understanding, is needed to make an informed decision whether or not to participate in a RCT (Kerr et al. 2004).

Interventions for control groups also raise ethical concerns. According to the principle ‘do not harm’, it is problematic to withhold therapies that are known to be effective. Hence, the control group should receive the **current standard of care** (Lo 2006). The control group shall not be denied a net superior medically established procedure for a specific condition that is under study for the population that the control group represents. One of the advantages of conceiving the standard of care in this way is that placebo-controlled trials may be compatible with equipoise (van der Graaf and van Delden 2011). There is also a growing interest in the ethics of using ‘active’ placebos in surgical trials, where otherwise there are no inert procedures available, and in pharmacological trials, where there are inert substances, but where patients may guess to which arm they have been allocated. An **inert placebo** is known not to have beneficial (or harmful) consequences caused by ‘active’ components or ingredients, and it may only superficially mimic the ‘active’ treatment by taste, shape, or colour, for example. **Active placebos** are designed and used to mimic some of the side effects of the intervention to further secure blinding. The root of the difference between the expected utility of inert and active placebo-controlled trials lies in the active placebo’s negative treatment effects, which may be, but are not always, minor. Using active placebos as controls when the situation is contrived so as to restrict access to a preferred intervention may be unacceptably exploitative. This is because participants have not had a free choice to avoid being controls who are offered care that is expected to be worse than care avail-

able routinely. Society as a whole and IRBs in particular may reasonably determine that setting up a situation where a group of patients are effectively forced to suffer a clear loss in prospect for the common good is unacceptable (Edward et al. 2005). Dilemmas about the control group are particularly difficult when the research participants have such a poor access to care that the research project is the only practical way for them to receive adequate health care (Lo 2006).

It is also unethical to continue a clinical trial if there is compelling evidence that one arm is safer or more effective. Furthermore, it would be wrong to continue a trial that will not answer to the research question because of low enrolment, few outcome events, or high dropout rates. The periodic analysis of interim data in a clinical trial by independent DMSBs (data safety monitoring boards) can determine if a trial should be terminated prematurely. Procedures for examining interim data and statistical stopping rules should be specified in the protocol (Lo 2006). Decisions by data safety and monitoring boards to stop trials early have potential implications for three distinct groups: research subjects currently enrolled in the trial, prospective subjects who may be recruited around or after the time of a decision to continue a trial after a board audit, and the population of patients, both current and future, who stand to benefit from clarification of an important clinical question. However, in research, the investigator’s principal ethical obligations are to design the study so as to minimize risk, to ensure the adequate disclosure of the remaining risks to prospective subjects, and to protect individual subjects enrolled in the study. This much is uncontroversial (Slutsky and Lavery 2004).

The European Medicines Agency (EMA), the British National Health Service Health Technology Assessment Program, and the United States Food and Drug Administration (FDA) have all recently published recommendations regarding the roles, responsibilities, and functions of DMSBs (EMA Committee for Medicinal Products for Human Use 2005; Grant et al. 2005; United States Department of Health and Human Services, Food and Drug Administration 2006). DMSBs are advisory boards which report to a

trial's sponsor and/or steering committee but work on behalf of trial participants and the broader public. Their principal mission is to monitor interim trial data for early evidence of significant harm or benefit. DSMBs members are typically individuals with clinical and/or clinical trials expertise. Members with ethics expertise and members of the lay public may also be helpful in some situations. The EMEA and FDA recommend DSMBs for randomized studies assessing interventions that are likely to be associated with important impacts on morbidity or mortality. Controversy exists regarding the threshold at which trials should be stopped for early evidence of benefit or harm, the extent of data sharing between DSMBs, and DSMB's current lack of transparency (Hicks et al. 2007).

2.3.2 Research on Previously Collected Specimens and Data

In reviewing the new developments in genetic science and medicine, and particularly the development of population genetics databases, it seems clear that the old notion of informed consent becomes outdated and needs review. Informed consent usually covers only specific and known uses of biological material—for example, when a person consents to their DNA being used in a particular experiment or research study. For other situations in which material may be used more than once, or for as yet unknown research, other forms of consent have been devised. Open, or blanket, consent is given only once but covers any use of the material at any time in the future. This is particularly important for scientific research, in which new projects or experiments might be devised years after individuals have given their consent and deposited their biological material; they may even have died in the meantime. Informed consent is given only after the patient or participant in a study has received information about the planned use of the material and health-care data. Presumed consent, conversely, assumes that an individual agrees in principle to their material being used for any reason: if not, they must withdraw their consent, or 'opt

out'. Presumed consent may be easier to obtain but understandably can alienate participants, who may resent their involuntary involvement. It is not clear yet whether the two main alternatives—presumed and open consent—will fare any better. Old rules often cannot fit new situations, and the changing needs, knowledge, and globalization in biomedical and genetic research may demand a new ethical and legal framework for consent (Kegley 2004).

Research on previously collected specimens and data offers no physical risks to participants. Consent for future studies is problematic because no one can anticipate what kind of research might be carried out later. Breaches of confidentiality may occur and lead stigma and discrimination (Lo 2006).

References

- Alderson P, Goodey C. Theories of consent. *BMJ*. 1998;317:1313–5.
- Barnett ML. Ethical issues in sponsored clinical research. *J Dent Res*. 1995;74:1129–32.
- Belmont Report: Office of Human Subjects Research. The Belmont report: ethical principles and guidelines for the protection of human subjects of research. 2011. Available at: <http://ohsr.od.nih.gov/guidelines/belmont.html>. Accessed 10 Dec 2011.
- Bhutta ZA. Ethics in international health research: a perspective from the developing world. *Bull World Health Organ*. 2002;80:114–20.
- Brody BA. Making informed consent meaningful. *IRB*. 2001;23(5):1–5. <http://www.ncbi.nlm.nih.gov/pubmed?term=Brody%20BA.%20Making%20informed%20consent%20meaningful>
- Brody BA, McCullough LB, Sharp RR. Consensus and controversy in clinical research ethics. *JAMA*. 2005;294:1411–4.
- Cahana A, Hurst SA. Voluntary informed consent in research and clinical care: an update. *Pain Pract*. 2008;8:446–51.
- Campbell FA, Goldman BD, Boccia ML, Skinner M. The effect of format modifications and reading comprehension on recall of informed consent information by low-income parents: a comparison of print, video, and computer-based presentations. *Patient Educ Couns*. 2004;53:205–16.
- Christiansen SL. Ethical and legal guidance in biomedical publishing: the AMA manual of style, tenth edition. Chest. 2008;134:1344–6.
- Council for International Organizations of Medical Sciences. International ethical guidelines for biomedical research involving human subjects. Geneva:

- Council for International Organizations of Medical Sciences; 1993.
- Dresden GM, Levitt MA. Modifying a standard industry clinical trial consent form improves patient information retention as part of the informed consent process. *Acad Emerg Med*. 2001;8:246–52.
- Edward SJ, Stevens AJ, Braunholtz DA, Lilford RJ, Swift T. The ethics of placebo-controlled trials: a comparison of inert and active placebo controls. *World J Surg*. 2005;29:610–4.
- EMA Committee for Medicinal Products for Human Use. Guideline on data monitoring committees. 2005. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003635.pdf. Accessed 10 Dec 2011
- Endacott R. Clinical research 2: legal and ethical issues in research. *Intensive Crit Care Nurs*. 2004;20(5):313–5.
- Erler CJ, Thompson CB. Part II: ethics, human rights, and clinical research. *Air Med J*. 2008;27:110–3.
- Ellen RP, Singleton R. Human rights and ethical considerations in oral health research. *J Can Dent Assoc*. 2008 Jun;74(5):439a–439d
- Faggion CM. Policies of dental journals for reporting and monitoring authorship and contributorship. *Br Dent J*. 2011;211:223–7.
- Falagas ME, Korbila IP, Giannopoulou KP, Kondilis BK, Peppas G. Informed consent: how much and what do patients understand? *Am J Surg*. 2009;198:420–35.
- Flanagin A, Carey LA, Fontanarosa PB, Phillips SG, Pace BP, Lundberg GD, Rennie D. Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *JAMA*. 1998;280:222–4.
- Flory J, Emanuel E. Interventions to improve research participants' understanding in informed consent for research: a systematic review. *JAMA*. 2004;292:1593–601.
- Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317:141–5.
- Grant AM, Altman DG, Babiker AB, Campbell MK, Clemens FJ, Darbyshire JH, Elbourne DR, McLeer SK, Parmar MKB, Pocock SJ, Spiegelhalter DJ, Sydes MR, Walker AF, Wallace SA, the DAMOCLES study group. Issues in data monitoring and interim analysis of trials. *Health Technol Assess*. 2005;9:1–235.
- Greco D, Diniz NM. Conflicts of interest in research involving human beings. *J Int Bioethique*. 2008;19(143–54):202–3.
- Hicks LK, Laupacis A, Slutsky AS. A primer on data safety monitoring boards: mission, methods, and controversies. *Intensive Care Med*. 2007;33:1815–8.
- Human D, Fluss SS. The World Medical Association's Declaration of Helsinki: historical and contemporary perspectives. 2001. Available at: http://www.wma.net/en/20activities/10ethics/10helsinki/draft_historical_contemporary_perspectives.pdf. Accessed 9 Dec 2011.
- Iverson C, Christiansen S, Flanagin A, et al. AMA manual of style: a guide for authors and editors. 10th ed. New York: Oxford University Press; 2007.
- Jeffcoat MK. A well-founded trust: a seven-tier defense against scientific misconduct. *J Am Dent Assoc*. 2002;133(7):804–6.
- Jefford M, Moore R. Improvement of informed consent and the quality of consent documents. *Lancet Oncol*. 2008;9(5):485–93.
- Jefford M, Mileskin L, Raunow H, et al. Satisfaction with the decision to participate in cancer clinical trials (CCT) is high, but understanding is a problem. *J Clin Oncol*. 2005;23:60–7.
- Kegley JA. Challenges to informed consent. *EMBO Rep*. 2004;5:832–6.
- Kerr C, Robinson E, Stevens A, Braunholtz D, Edwards S, Lilford R. Randomisation in trials: do potential trial participants understand it and find it acceptable? *J Med Ethics*. 2004;30:80–4.
- Kim SY, Caine ED, Currier GW, Leibovici A, Ryan JM. Assessing the competence of persons with Alzheimer's disease in providing informed consent for participation in research. *Am J Psychiatry*. 2001;158:712–7.
- Lo B. Addressing ethical issues. In: Hulley SB, Cummings SR, Browner WS, Grady DG, Thomas B, Newman TB, editors. *Designing clinical research*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2006. p. 225–39.
- Macklin R. Understanding informed consent. *Acta Oncol*. 1999;38:83–7.
- Mann H. Extensions and refinements of the equipoise concept in international clinical research: would Benjamin Freedman approve? *Am J Bioeth*. 2006;6:67–9.
- Michael K, Paasche-Orlow, Holly A, Taylor, Frederick L. Brancati, Readability Standards for Informed-Consent Forms as Compared with Actual Readability. *NEJM*. 2003;348:721–726.
- Morin K, Rakatansky H, Riddick Jr FA, Morse LJ, O'Bannon 3rd JM, Goldrich MS, Ray P, Weiss M, Sade RM, Spillman MA. Managing conflicts of interest in the conduct of clinical trials. *JAMA*. 2002;287:78–84.
- Mowatt G, Shirran L, Grimshaw JM, Rennie D, Flanagin A, Yank V, MacLennan G, Göttsche PC, Bero LA. Prevalence of honorary and ghost authorship in Cochrane reviews. *JAMA*. 2002;287:2769–71.
- Nicholl J. The ethics of research ethics committees. *BMJ*. 2000;320(7243).
- Office for Protection from Research Risks. Protection of human subjects, 45 CFR 46. Office for Protection from Research Risks Reports, Federal Register. 1996.
- Paasche-Orlow MK, Paasche-Orlow MK, Taylor HA, Brancati FL. Readability standards for informed-consent forms as compared with actual readability. *N Engl J Med*. 2003;348:721–6.
- Pandya A. Readability and comprehensibility of informed consent forms for clinical trials. *Perspect Clin Res*. 2010;1:98–100.
- Pitak-Arnnop P, Sader R, Hervé C, Dhanuthai K, Bertrand JC, Hemprich A. Reporting of ethical protection in recent oral and maxillofacial surgery research involving human subjects. *Int J Oral Maxillofac Surg*. 2009;38:707–12.

- Responsibilities of the Office of Human Subjects Research (OHSR). 2011. Available at: <http://ohsr.od.nih.gov/info/sheet1.html>. Accessed 10 Dec 2011.
- Rice TW. The historical, ethical, and legal background of human-subjects research. *Respir Care*. 2008;53:1325–9.
- Rivera R, Borasky D, Rice R, Carayon F, Wong E. Informed consent: an international researchers' perspective. *Am J Public Health*. 2007;97:25–30.
- Robinson EJ, Kerr C, Stevens A, Lilford R, Brauholtz D, Edwards S. Lay conceptions of the ethical and scientific justifications for random allocation in clinical trials. *Soc Sci Med*. 2004;58:811–24.
- Robinson EJ, Kerr CE, Stevens AJ, Lilford RJ, Brauholtz DA, Edwards SJ, Beck SR, Rowley MG. Lay public's understanding of equipoise and randomisation in randomised controlled trials. *Health Technol Assess*. 2005;9:1–192.
- Sekine T, Shimada M. Protection of human subjects in medical research: from the viewpoint of historical development of ethical regulations. *Nihon Ishigaku Zasshi*. 2011;57:63–70.
- Silverman HJ, Luce JM, Lanken PN, et al. Recommendations for informed consent forms for critical care clinical trials. *Crit Care Med*. 2005;33:867–82.
- Slutsky AS, Lavery JV. Data safety and monitoring boards. *N Engl J Med*. 2004;350:1143–7.
- Smith R. The trouble with medical journals. *J R Soc Med*. 2006;99:115–9.
- United States Department of Health and Human Services, Food and Drug Administration. 2006. Guidance for clinical trial sponsors. Establishment and operation of clinical trial data monitoring committees. Available at: <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM127073.pdf>. Accessed 10 Dec 2011.
- van der Graaf R, van Delden JJ. Equipoise should be amended, not abandoned. *Clin Trials*. 2011;8:408–16.
- Verástegui EL. Consenting of the vulnerable: the informed consent procedure in advanced cancer patients in Mexico. *BMC Med Ethics*. 2006;7:E13.
- Wislar JS, Flanagan A, Fontanarosa PB, Deangelis CD. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. *BMJ*. 2011;343:d6128.
- Wolf LE, Lo B. Association of Professors of Medicine. Ethical issues in clinical research: an issue for all internists. *Am J Med*. 2000;109:82–5.
- World Medical Association declaration of Helsinki. Recommendations guiding physicians in biomedical research involving human subjects. *JAMA*. 1997;277:925–6.
- World Medical Association. Declaration of Helsinki. 2011. Available at: <http://www.wma.net/en/60about/70history/01declarationHelsinki/>. Accessed 9 Dec 2011.

The practice of periodontology continues to increase in complexity. Developments in therapies and techniques, changing socio-demographic patterns, increasingly knowledgeable health care consumers, and the information ‘explosion’ all are placing greater demands on clinical decision making. As health care practitioners, it is important to offer the best possible care for patients (Worthington and Needleman 2005).

1.1 What Is Evidence-Based Medicine?

Users of medical research, especially clinical practitioners, focus primarily on accessing results of clinical studies to answer clinical questions, most frequently, ‘does this intervention work?’ Comparatively little attention is paid to the questions themselves. Yet failure to think carefully about the meaning, structure, and intention of research questions can have adverse effects on every subsequent step of the research process, potentially compromising the answers. The fundamental purpose of asking good questions is to match these to an appropriate and feasible study design (Bragge 2010).

Evidence-based medicine (EBM) is ‘the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients’ (Sackett et al. 1996; Haroon and Phillips 2010). It helps support the practice of making clinical decisions which are informed by the best evidence ‘out there’ and by a patient’s beliefs and values. It also offers the potential to partially solve one of the big issues of the

information age, which is how to handle an exponential increase in the amount of medical information ‘out there’ without inventing a 28-h day (Haroon and Phillips 2010; Schardt et al. 2007).

EBM consists of five separate stages (Sackett, 2000); Scherer and Smith, 2002 Phillips and Glasziou 2008; Haroon and Phillips 2010; Tanjong-Ghogomu et al. 2009).

1.1.1 Asking a Clinical Question

Clinical questions are generated from clinical problems and are usually related to diagnosis, aetiology, prognosis, or treatment or prevention. The importance of a question varies according to the perspective of the person asking it. The introduction of new treatments leads to questions such as those related to efficacy and safety. A well-built clinical question should have four components that can be identified by the acronym PICO (Armstrong 1999; Bergus and Emerson 2005; Booth, 2006; Larue et al. 2009; Santos et al. 2007; Oxman et al. 1994; Counsell 1997; Richardson et al. 1995; Tanjong-Ghogomu et al. 2009):

- **P** for patient population: a description of the patient or the target disorder of interest
- **I** for intervention: could include an exposure, a diagnostic test, a prognostic factor, a therapy, or a patient perception, etc.
- **C** for comparison intervention: could be a placebo or active treatment
- **O** for outcome: a clinical outcome of interest to you and your patient

As an example, a PICO question was developed from a hypothetical case scenario (Faggion and Tu 2007). A 56-year-old, systematically healthy male patient with chronic periodontitis is looking for treatment. He is afraid of dental procedures and would like to avoid staying long in a dental chair. This information was converted into one well-structured question in the PICO format, which was used to direct the literature searching:

P (*patient*): adults with chronic periodontitis

I (*intervention*): periodontal treatment with sonic and ultrasonic scalers

C (*comparison*): periodontal treatment with manual instruments

O (*outcome*): effectiveness (measured by clinical attachment gain), pocket probing depth reduction, and efficiency (measured by treatment time)

Therefore, the question was structured as follows: 'In adult patients with chronic periodontitis, which nonsurgical periodontal treatment (sonic/ultrasonic scalers or manual instruments) is more effective, in terms of probing pocket depth reduction/attachment level gain, and efficient in terms of duration of treatment?' (Faggion and Tu 2007).

Different clinical research questions require evaluation through different study designs. A study to determine the effectiveness of surgical therapy compared with nonsurgical debridement deals with the effectiveness of a treatment option and would be best answered by a randomized controlled trial (RCT) or, ideally, a systematic review of RCTs. However, it must be noted that although RCTs and systematic reviews of RCTs may well be the 'gold standard' upon which to base decisions on the effectiveness of interventions, they are not necessarily appropriate, or ethical, to answer all questions. An RCT would obviously not be helpful in answering the question posed on the epidemiological evidence of plaque in the aetiology of periodontitis. For such questions regarding prognosis or aetiology, cohort studies would be more appropriate. **Table 1.1** illustrates the types of study designs most suitable

for different types of research questions arising in periodontology. The most appropriate source of information will depend upon the type of study design being sought (Needleman et al. 2005).

1.1.2 Locating the Evidence

The next step in EBM methods is searching for the best evidence available. Clinicians have the responsibility of seeking the best evidence to guide their practice. There are two sources of evidence: primary and secondary sources. **Primary literature** can be searched in Google and on electronic databases, such as MEDLINE and EMBASE, for relevant journals, reference lists, and conference proceedings. Similarly, printed materials can be accessed in libraries from indexed sources; drug companies may be contacted; grey literature can be searched; and personal communications can be sent. However, there are over 18 million journal citations in MEDLINE, which is an unmanageable amount of information to explore (Wilczynski et al. 2002). **Secondary or pre-appraised literature**, such as systematic reviews, helps make sense of unmanageable amounts of research, bringing together separately conducted studies and synthesizing their results. The Cochrane Library is an outstanding source of the most reliable, independent evidence on which to base clinical treatment decisions, as Cochrane reviews are produced to the highest methodological standard. Cochrane reviews are done on a variety of health care diseases and conditions, including periodontal disease (Cook et al. 1997; Tanjong-Ghogomu et al. 2009).

1.1.3 Appraising and Synthesizing the Evidence

The appropriate information from the search is then selected for critical appraisal to evaluate the validity and usefulness of the evidence. Good medical research must have sufficient

Table 1.1 Study designs and the types of questions they address (Needleman et al.2005) (Reprinted with permission from John Wiley & Sons)

Definition of study design	Used for (examples given in <i>italics</i>)
<i>Experimental studies</i>	
Randomized controlled trial: parallel group design – a group of participants (or other unit of analysis, e.g. teeth) is randomized into different treatment groups. These groups are followed up for the outcomes of interest	Evaluating the effectiveness of an intervention <i>Randomized controlled trial comparing the effectiveness of surgical therapy and nonsurgical debridement</i>
Randomized controlled trial: split-mouth design – each patient is his/her own control. A pair of similar teeth, or groups of teeth (quadrants), may be selected and randomly allocated to different treatment groups	
Non-randomized controlled trial – allocation of participants under the control of the investigator, but the method falls short of genuine randomization	<i>Controlled trial comparing two methods of treating periodontal intrabony defects using pairs of sites where the LHS is always group A and the RHS group B</i>
<i>Observational studies</i>	
Cohort: a longitudinal study, identifying groups of participants according to their exposure/intervention status. Groups are followed forward in time to measure the development of different outcomes	Measuring the incidence of a disease; looking at the causes of disease; determining prognosis. Cohort study looking at the progress of periodontitis over time and relating this to external factors such as smoking or plaque
Case-control: a study that identifies groups of participants according to their disease/outcome status. Groups are investigated/questioned to determine their exposure status	Identifying potential risk factors for a disease; looking at the possible causes of disease. <i>Case-control study looking at the prevalence of periodontitis and relating this to factors such as genetic markers</i>
Cross-sectional: a study (survey) undertaken on a defined population at a single point in time (snapshot). Subjects are observed on just one occasion and are not followed up	Measuring the prevalence of a disease or risk factor in a defined population at a specific time. <i>A cross-sectional study to determine the current periodontal treatment needs in a specific population</i>

methodological rigour to be reliable enough to answer clinical questions (Rosenberg and Donald 1995; Oxman et al. 1994; Tanjong-Ghagomu et al. 2009).

So how can we assess whether a paper is worth reading? One method of achieving this is to have a systematic process of assessing the reliability, relevance, and results of published papers. This is referred to as **critical appraisal or critical reading**. Critical appraisal can be criticized by anyone with access to appropriate methodologically sound appraisal criteria. Appraisal criteria can be found at several websites (ADA Center for Evidence-Based Dentistry <http://ebd.ada.org/Resources.aspx>; Centre for Evidence-Based Medicine, UK <http://www.cebm.net/?o=1040>), and there are also several books and CD-ROMs that now give information on how to carry this out (Richards 2000).

However, despite that range of criteria available for appraisal of papers, there are only three essential questions that need to be asked of any paper, as proposed by Richards (2000):

1. Is the study valid?

By valid we mean was the study conducted properly. This is also referred to as a study's internal validity, hence the question. To assess a study's internal validity, we can ask a number of subsidiary questions: Did the authors ask a clear question? Did they use the most appropriate study design? Was the study carried out reliably? Are the methods clearly described so others can see what has been done? (Richards 2000).

2. What are the results?

If the study has been carried out properly, you can look to the results. We can then ask: Is there a clear result? Are the results presented clearly using appropriate statistical tests? (Richards 2000).

3. Are the results relevant?

Once you have appraised the internal validity and the results of the paper, you can ask if they are relevant. What we mean here is, are the results relevant to my patients? Are the group of patients on whom the research was carried out similar enough to the ones I see every day to apply the findings of the paper to them? (Richards 2000).

In terms of **strength of evidence**, it is the validity of a piece of evidence that is important. The validity of a study is the extent to which its design and conduct are likely to prevent systematic errors or bias. Therefore, the more valid a piece of evidence, the greater its strength and the more secure you can feel making treatment decisions based on it. The need to develop a method of ranking the validity of evidence was initially developed by Fletcher and Sackett while working on the Canadian Task Force on Periodic Health Examination (Canadian Task Force on the Periodic Health Examination 1979). These initial levels and grades were criticized, however, for their therapeutic/therapeutic/preventive orientation. Consequently, the need to develop similar levels for diagnostic, prognostic, harm, and economic studies led a group of people associated with the Centre for Evidence-Based Medicine to develop a more complete level-of-evidence table (Table 1.2), with associated grades of recommendation. This version printed below from 2005 was last modified in 2011, and it is also available on the Centre for Evidence-Based Medicine website (www.cebm.net) where it is under constant review, so readers should visit the site from time to time to check for changes (Richards 2003).

1.1.4 Applying the Evidence

The conclusions of this search and critical appraisal are worthwhile only if they are translated into actions that affect our patients. To apply evidence to a patient clinical management,

it is critical to discuss the evidence, the benefits and harms, and the alternative treatments with patients so that they understand. Patient decision aids present the treatments as a choice of options and include personalized information about options, outcomes, probabilities, and uncertainties in sufficient detail for decision-making. Decision aids improve people's knowledge of the options, create accurate risk perceptions of their benefits and harms, reduce difficulty with decision-making, and increase participation in the process (Evidence-Based Medicine Working Group 1992; Sackett and Rosenberg 1995; O'Connor 2001; Tanjong-Ghagomu et al. 2009).

1.1.5 Self-Evaluation

The fifth step in practicing EBM is self-evaluation. It involves evaluating your performance in:

- Asking answerable clinical questions
- Finding the best evidence
- Critically evaluating the evidence for its validity and potential usefulness
- Integrating the critical appraisal with clinical expertise and applying the results in clinical practice

Another step in the self-evaluation process is considering your patient's reactions or improved health care. This will be influenced by their participation in the decision-making process, their adherence to treatment, any adverse events they experience, and their general health status. The practice of the five steps of EBM is a way forward to improve the incorporation of best evidence into clinical practice, as well as to maintain and expand clinically important knowledge. However, EBM principles and practice are not yet sufficiently disseminated, and there is also the challenge for clinicians to integrate it into routine clinical practice (Evidence-Based Medicine Working Group 1992; Tanjong-Ghagomu et al. 2009).

Table 1.2 Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence (OCEBM Levels of Evidence Working Group.^a “The Oxford 2011 Levels of Evidence”) (<http://www.2.cch.org.tw/ebm/file/CEBM-Levels-of-Evidence.pdf>) (Reprinted with permission)

Question	Step 1 (Level 1b)	Step 2 (Level 2b)	Step 3 (Level 3b)	Step 4 (Level 4b)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances	Local non-random sample	Case-series	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standard	Case-control studies, or poor or non-independent reference standard	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial	Case-series or case-control studies, or poor quality prognostic cohort study	n/a
Does this intervention help? (treatment benefits)	Systematic review of randomized trials or <i>n-of-1</i> trials	Randomized trial or observational study with dramatic effect	Non-randomized cohort/follow-up study	Case-series, case-control studies, or historically controlled studies	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, <i>n-of-1</i> trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.) ^c	Case-series, case-control, or historically controlled studies	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or <i>n-of-1</i> trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non-randomized cohort/follow-up study	Case-series, case-control, or historically controlled studies	Mechanism-based reasoning

Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653>

^aOCEBM Table of Evidence Working Group = Jeremy Howick, Iain Chalmers (James Lind Library), Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard and Mary Hodgkinson

^bLevel may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size

^cAs always, a systematic review is generally better than an individual study

1.2 What Is Evidence-Based Dentistry?

The CDA Task Force on Evidence-Based Dentistry recommended a definition of EBD drawn from the ‘Oral Health in America’ report by the US Surgeon General, which is philosophically consistent with the American Dental Association’s definition (Kao 2006).

The American Dental Association defines the term ‘evidence-based dentistry’ as follows: ‘Evidence-based dentistry (EBD) is an approach to oral health care that requires the judicious integration of systematic assessments of clinically relevant scientific evidence, relating to the patient’s oral and medical condition and history, with the dentist’s clinical expertise and the patient’s treatment needs and preferences’. This definition stresses the importance of three elements: a dentist’s expertise and clinical judgement, relevant clinical evidence that is present in the literature, and the informed patient’s preference. In a dental practice that incorporates an evidence-based approach, the practitioner’s experience is primary since it is his or her responsibility to consider all three components when defining the best course of treatment. Ideally, evidence-based treatment is characterized by the intersection of these three elements (Kao 2006).

1.3 What Is Evidence-Based Periodontology?

Evidence-based periodontology is the application of evidence-based health care to periodontology and can be considered a tool to support decision-making and integrating the best evidence available with clinical practice (Needleman et al. 2005). The findings of research need to be translated into information that is useful for each health decision-maker, including the patient. Using a standard, unbiased method to evaluate information is considered better and necessary because the number of new scientific insights that emerge each year is overwhelming (Newman et al. 2003).

The highest quality evidence will be used if it exists, but if it does not, lower levels of evidence will be considered. Lower levels of evidence usually means research designs more prone to bias and therefore with less reliable data. However, the nature, strengths, and weaknesses of the evidence will be made clear to the reader. In addition, wherever possible, the data presentation supplies more clinically relevant information, including the probability of achieving a certain effect such as a benefit, and considering possible adverse effects (Needleman et al. 2005).

1.4 Evidence-Based Periodontology Versus Traditional Periodontology

High-quality research and the use of evidence are fundamental to both evidence-based periodontology and traditional periodontology. The differences between these approaches emanate from how research informs clinical practice. Evidence-based periodontology uses a more transparent approach to acknowledge both the strengths and the limitations of the evidence. An appreciation of the level of uncertainty or imprecision of the data is essential in order to offer choices to the patient regarding treatment options. Evidence-based periodontology also attempts to gather all available data and to minimize bias in summarizing the data. Furthermore, evidence-based periodontology acknowledges explicitly the type or level of research on which conclusions are drawn. In contrast, traditional periodontology uses unclear basis of evidence and unclear or absent appraisal of quality of evidence; presets a more subjective, more opaque, and more biased process; and shows a greater tendency to black and white conclusions (Needleman et al. 2005).

References

- American Dental Association, ADA Positions & Statements. ADA policy on evidence-based dentistry. 2011. Available at: <http://www.ada.org/1754.aspx>. Accessed 2011.

- Armstrong EC. The well-built clinical question: the key to finding the best evidence efficiently. *WMJ*. 1999;98:25–8.
- Bergus GR, Emerson M. Family medicine residents do not ask better-formulated clinical questions as they advance in their training. *Fam Med*. 2005;37:486–90.
- Booth A. Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*. 2006;24:355–68.
- Bragge P. Asking good clinical research questions and choosing the right study design injury. *Injury*. 2010;41 Suppl 1:S3–6.
- Canadian Task Force on the Periodic Health Examination. The periodic health examination. *Can Med Assoc J*. 1979;121:1193–254.
- Cook DJ, Greengold NL, Ellrodt AG, Weingarten SR. The relation between systematic reviews and practice guidelines. *Ann Intern Med*. 1997;127:210–6.
- Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews. *Ann Intern Med*. 1997;127:380–7.
- Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*. 1992;268:2420–5.
- Faggion Jr CM, Tu YK. Evidence-based dentistry: a model for clinical practice. *J Dent Educ*. 2007;71:825–31.
- Haroon M, Phillips R. “There is nothing like looking, if you want to find something” – asking questions and searching for answers – the evidence based approach. *Arch Dis Child Educ Pract Ed*. 2010;95:34–9.
- Kao RT. The challenges of transferring evidence-based dentistry into practice. *J Evid Based Dent Pract*. 2006;6:125–8.
- Larue EM, Draus P, Klem ML. A description of a Web-based educational tool for understanding the PICO framework in evidence-based practice with a citation ranking system. *Comput Inform Nurs*. 2009;27:44–9.
- Needleman I, Moles DR, Worthington H. Evidence-based periodontology, systematic reviews and research quality. *Periodontol* 2000. 2005;37:12–28.
- Newman MG, Caton JG, Gunsolley JC. The use of the evidence-based approach in a periodontal therapy contemporary science workshop. *Ann Periodontol*. 2003;8:1–11.
- O'Connor A. Using patient decision aids to promote evidence-based decision making. *Evid Based Med*. 2001;6:100–2.
- Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653> Accessed 2011.
- Oxman AD, Cooc DJ, Guyatt GH. Users' Guides to the medical literature. VI. How to use an overview. Evidence-based medicine working group. *JAMA*. 1994;272:1367–71.
- Phillips R, Glasziou P. Evidence based practice: the practicalities of keeping abreast of clinical evidence while in training. *Postgrad Med J*. 2008;84:450–3.
- Richards D. Is it worth reading this paper? *Evid Based Dent*. 2000;2:50–2.
- Richards D. Not all evidence is created equal – so what is good evidence? *Evid Based Dent*. 2003;4:17–8.
- Richardson W, Wilson M, Nishikawa J, Hayward R. The well-built clinical question: a key to evidence-based decisions. *ACP J Club*. 1995;123:A12–3.
- Rosenberg W, Donald A. Evidence based medicine: an approach to clinical teaching. *BMJ*. 1995;310:1122–6.
- Sackett DL. Evidence-based medicine: how to practice and teach EBM. 2nd ed. Edinburgh, New York: Churchill Livingstone; 2000. p. xiv, 261.
- Sackett DL, Rosenberg WMC. On the need for evidence-based medicine. *J Public Health Med*. 1995;17(3):330–4.
- Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312:71–2.
- Santos CMC, Pimenta CAM, Nobre MRC. The PICO strategy for the research question construction and evidence search. *Rev Lat Am Enfermagem*. 2007;15:508–11.
- Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak*. 2007;7:16.
- Scherer S, Smith MB. Teaching evidence-based practice in academic and clinical settings. *Cardiopulm Phys Ther*. 2002;13:23.
- Tanjong-Ghogu E, Tugwell P, Welch V. Evidence-based medicine and the Cochrane Collaboration. *Bull NYU Hosp Jt Dis*. 2009;67:198–205.
- Wilczynski NL, Haynes RB, Hedges Team. Robustness of empirical search strategies for clinical content in MEDLINE. *Proc AMIA Annu Symp*. 2002:904–8.
- Worthington H, Needleman I. Evidence-based periodontal disease prevention and treatment: introduction. *Periodontol* 2000. 2005;37:9–11.

Reading research is mandatory if a clinician is to keep up. Analogous to biological taxonomy, a simple hierarchy can be used to categorize most studies. To do so, however, the study design must be known. As in biology, anatomy dictates physiology. The anatomy of a study determines what it can and cannot do. A difficulty that readers encounter is that authors sometimes do not report the study type or provide sufficient detail to figure it out. A related problem is that authors sometimes incorrectly label the type of research done (Grimes and Schulz 2002a).

Different types of clinical trials can be recognized (ClinicalTrials.gov 2011):

- **Treatment trials** test experimental treatments, new combinations of drugs, or new approaches to surgery or radiation therapy.
- **Prevention trials** look for better ways to prevent disease in people who have never had the disease or to prevent a disease from returning. These approaches may include medicines, vaccines, vitamins, minerals, or lifestyle changes.
- **Diagnostic trials** are conducted to find better tests or procedures for diagnosing a particular disease or condition.
- **Screening trials** test the best way to detect certain diseases or health conditions.
- **Quality of life trials** (or supportive care trials) explore ways to improve comfort and the quality of life for individuals with a chronic illness.

Clinical trials are conducted in phases. The trials at each phase have a different purpose and help scientists answer different questions (ClinicalTrials.gov 2011):

- **In phase I trials:** Initial studies to determine the metabolism and pharmacologic actions of drugs in a small group of people (20–80), the side effects associated with increasing doses, and to gain early evidence of effectiveness; these may include healthy participants and/or patients.
- **In phase II trials:** Controlled clinical studies conducted to evaluate the effectiveness of the drug for a particular indication or indications in a larger group of people (100–300), patients with the disease or condition under study, and to determine the common short-term side effects and risks.
- **In phase III trials:** The experimental study drug or treatment is given to large groups of people (1,000–3,000) to confirm its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow the experimental drug or treatment to be used safely.
- **In phase IV trials:** Post-marketing studies delineate additional information including the drug's risks, benefits, and optimal use.

3.1 Developing a Research Question

Research is a process of asking questions and using highly structured techniques to answer these questions. Identifying the right question to address is an important first step in the research process (see Fig. 3.1). Different ways of devising and revising a research question, commencing with an initial idea

Fig. 3.1 A summary of the research process (Newton et al. 2004. Reprinted with permission from Macmillan Publishers Ltd.)

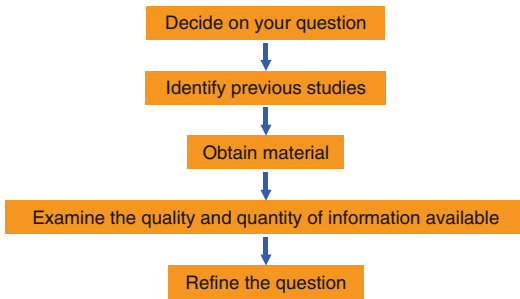
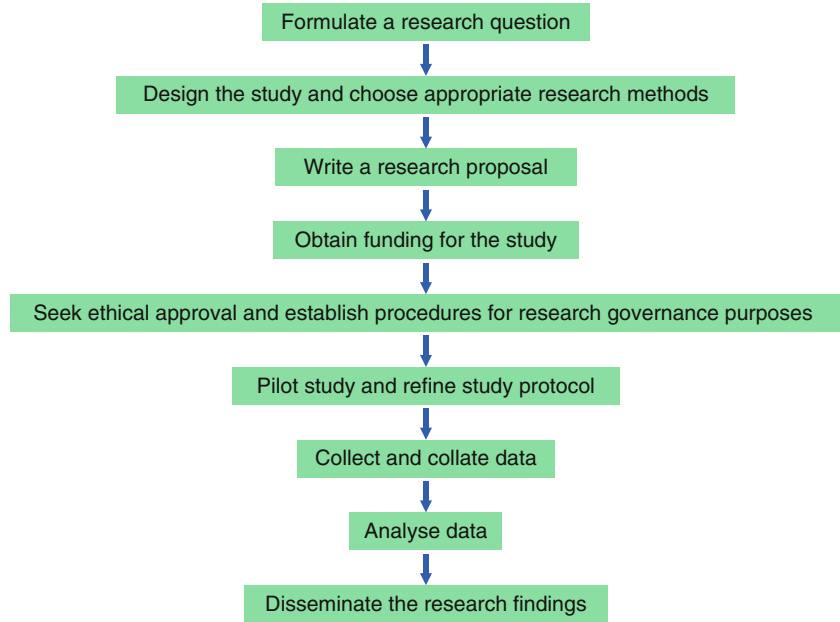


Fig. 3.2 Devising and revising your research question (Newton et al. 2004. Reprinted with permission from Macmillan Publishers Ltd.)

and focusing this into a testable research question, are described in **Fig. 3.2** (Newton et al. 2004).

3.2 Classification of Types of Clinical Research

So you have devised your research question, *how* are you going to answer it? Study designs can be divided into two broad categories: qualitative studies and quantitative studies. A study may be purely qualitative, purely quantitative, or combine elements of both designs (Williams et al. 2004).

Qualitative studies, for example, in-depth interviews and focus group work, aim to explore and obtain insight into complex issues, such as the

reasons for people's attitudes and behaviours. The results are described in words rather than numbers. Rather than aspiring to representativeness, qualitative research usually aims to reflect the *diversity* within a given population (Williams et al. 2004).

Quantitative studies aim to test a hypothesis. The results are given in numbers or proportions, which most general dental practitioners are familiar with and find easy to conceptualize. The quantitative approach aims to generate data which are *representative* of the given population. A brief introduction to some of the important aspects of the quantitative study design process is now given, and a summary is outlined in **Fig. 3.3** (Williams et al. 2004).

Quantitative study designs may be used to describe how often an event occurs (descriptive study), to examine the relationships between different phenomena (association study), or to compare two or more sample groups (comparative study) (Williams et al. 2004). Quantitative studies may be *experimental* or *non-experimental* (observational) (**Fig. 3.4**). For experimental studies, one needs to distinguish whether the exposures were assigned by a truly random technique (with concealment of the upcoming assignment from those involved) or whether some other allocation scheme was used, such as alternate assignment. With observational studies, which dominate the literature, the next step is to ascertain whether the study has a comparison

- Clarify aims (based on the research question) and formulate objectives
- State hypothesis (null hypothesis for comparative studies)
- Decide type of study eg cross-sectional survey, randomised controlled trial
- Give details of study intervention eg treatment, investigation, interview
- What are the dependent/independent/confounding variables?
- What is the study population from which samples will be drawn eg patients from your practice or other dental practices?
- Decide sampling methods and inclusion/exclusion criteria
- Decide sample size and justify it with a power calculation
- In a comparative study, how many groups will be studied, is a control group required and how will you allocate participants to each group?
- How will participants be recruited to the study?
- Will the participants/investigators be 'blind'?
- Choose appropriate instruments to measure outcomes
- Decide end-points for clinical trials
- Plan procedures eg order, site, timing, frequency, information given, equipment used, storage of samples etc.
- Consider how to reduce bias
- Consider how to reduce hazards and risks and deal with potential problems
- Plan method of data entry and the data analysis package to be used
- Plan analysis of the data, statistical tests to be used, level of significance etc

Fig. 3.3 Quantitative study design stages (Williams et al. 2004) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

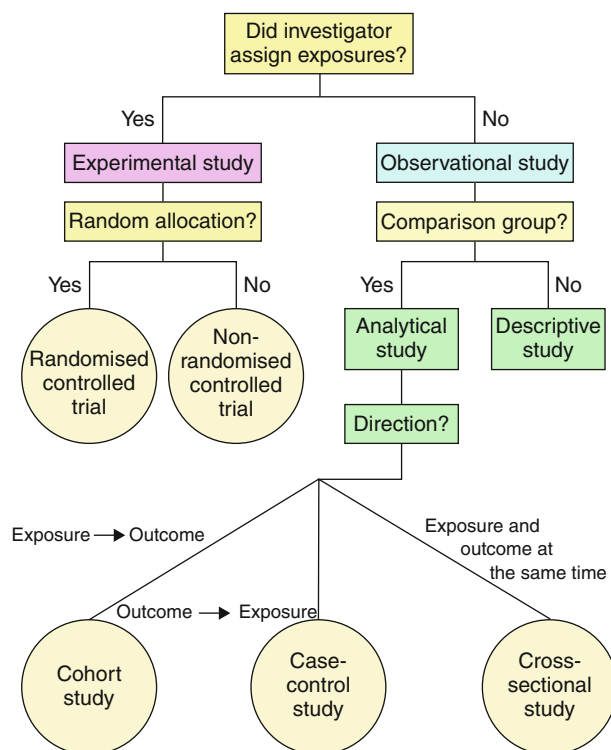


Fig. 3.4 Algorithm for classification of types of clinical research (Reprinted from Grimes and Schulz (2002a) *Lancet* 2002;359(9300):57–61, with permission from Elsevier)

or control group. If so, the study is termed analytical. If not, it is a descriptive study. If the study is analytical, the temporal direction of the trial needs to be identified. If the study determines both exposures and outcomes at one time point, it is termed cross-sectional. This type of study provides a snap-

shot of the population of sick and well at one time point. If the study begins with an exposure, then it is deemed a cohort study. Cohort studies can be either concurrent or non-concurrent. By contrast, if the analytical study begins with an outcome and looks back in time for an exposure, then the study

Table 3.1 What to look for in observational studies (Grimes and Schulz 2002e) Reprinted from The Lancet 359(9302):248–52 with permission from Elsevier

Is selection bias present?
• In a cohort study, are participants in the exposed and unexposed groups similar in all important respects except for the exposure? • In a case–control study, are cases and controls similar in all important respects except for the disease in question?
Is information bias present?
• In a cohort study, is information about outcome obtained in the same way for those exposed and unexposed? • In a case–control study, is information about exposure gathered in the same way for cases and controls?
Is confounding present?
• Could the results be accounted for by the presence of a factor – e.g., age, smoking, sexual behaviour, diet - associated with both the exposure and the outcome but not directly involved in the causal pathway?
If the results cannot be explained by these three biases, could they be the result of chance?
• What are the relative risk or odds ratio and 95% CI? • Is the difference statistically significant, and, if not, did the study have adequate power to find a clinically important difference?
If the results still cannot be explained away, then (and only then) might the findings be real and worthy of note.

is a case–control study. Studies without comparison groups are called descriptive studies. At the bottom of the research hierarchy is the case report. When more than one patient is described, it becomes a case-series report (Grimes and Schulz 2002a).

3.3 Observational Clinical Research

Observational clinical research poses numerous challenges. Among the most common challenges are those of **data quality** and access to data. Data quality results from a combination of precision (reproducibility) and accuracy (validity). Data collected for clinical purposes and reported in medical records may meet both, one, or neither of the requirements. Because much clinical research is performed retrospectively, data available in medical records may be missing, poorly collected, or erroneous. Laboratory tests that are adequate for clinical purposes may not have the precision and accuracy needed for research purposes. Observational clinical research is particularly susceptible to poor data quality unless great care is taken to design a study prospectively and ensure validation (Luepker 2005).

With respect to internal validity, selection **bias**, information bias, and confounding are present to some degree in all observational research. Selection bias stems from an absence of comparability

between groups being studied. Information bias results from incorrect determination of exposure, outcome, or both. The effect of information bias depends on its type. If information is gathered differently for one group than for another, bias results. By contrast, non-differential misclassification tends to obscure real differences. The challenge for investigators, editors, and readers is to ferret these out and judge how they might have affected results. A simple checklist, such as that shown in **Table 3.1**, can be helpful (Grimes and Schulz 2002e).

A descriptive study is ‘concerned with and designed only to describe the existing distribution of variables, without regard to causal or other hypotheses’ (Grimes and Schulz 2002b).

Good descriptive research, like good newspaper reporting, should answer **five basic ‘W’ questions**—who, what, why, when, and where—and an implicit sixth question, so what? So what? The implicit ‘W’ relates to the public health effect. In view of the proliferation of descriptive reports, what is their import? Is the condition a current and timely one? Is it serious? Are large numbers involved? Are its societal implications broad? Has it been studied before? Although many descriptive reports herald new illnesses or monitor health, the net effect of others might be only thicker curricula vitae at the expense of thinner forests (Grimes and Schulz 2002b) (**Table 3.2**).

Descriptive studies consist of two major groups: those that deal with individuals and those

Table 3.2 The descriptive pentad (Grimes and Schulz (2002b) (Reprinted from The Lancet 2002;359(9301):145-9 with permission from Elsevier)

Nr.	Question	Discussion
1	Who has the disease in question?	Age and sex are universally described, but other characteristics might be important too, including race, occupation, or recreational activities.
2	What is the condition or disease being studied?	Development of a clear, specific, and measurable case definition is an essential step in descriptive epidemiology. Without such a description, the reader cannot interpret the report.
3	Why did the condition or disease arise?	Descriptive studies often provide clues about cause that can be pursued with more sophisticated research designs.
4	When is the condition common or rare?	Time provides important clues about health events.
5	Where does or does not the disease or condition arise?	Geography has had a huge effect on health, such as presence of fluoridated water etc.

that relate to populations. Studies that involve individuals are the case report, the case-series report, cross-sectional studies, and surveillance, whereas ecological correlational studies examine populations (Grimes and Schulz 2002b).

Descriptive studies have several useful roles: (1) *Trend analysis*: Being able to monitor the health of populations is important to health-care administrators. Trend analysis is often provided by ongoing surveillance. (2) *Planning*: A second use is health-care planning. (3) A third use of descriptive studies is to develop *hypotheses about cause* (Grimes and Schulz 2002b).

Descriptive studies have both strengths and weaknesses. Often, the data are already available and thus inexpensive and efficient to use. Furthermore, few ethical difficulties exist. However, descriptive studies have important limitations. Temporal associations between putative causes and effects might be unclear. A dangerous pitfall is that the investigators might draw causal inferences when none is possible (Grimes and Schulz 2002b).

3.3.1 Case Report

The case report is the least publishable unit in the medical literature. Often, an observant clinician reports an unusual disease or association, which prompts further investigations with more rigorous study designs (Grimes and Schulz 2002b).

Currently, three basic types of case reports merit publication. The first is the ‘highly unique case’ that may represent a previously undescribed syndrome or disease. The second type is the case

that demonstrates an unexpected association between two or more diseases or disease manifestations and may represent ‘an unsuspected causal relationship’. The third type is the case with an unexpected outcome, suggesting a surprising ‘new beneficial therapeutic effect or adverse effect’. Case reports are often intended to be entertaining as well as educational (Thompson and Panacek 2007).

When the findings in a single case report are not fully adequate or appropriately representative of the disease process, including a discussion of previously described cases that have some features in common is typical. This kind of paper that collates and interprets previous reports, as well as the reported case, is often referred to as ‘a case report with a review of the literature’. Even in such situations, the reported case must be sufficient to stand largely on its own. If not, then it is best to drop the case entirely and consider writing an article that reviews the literature on the subject. In other words, a weak case report cannot be rescued simply by adding in a review of the literature (Thompson and Panacek 2007).

Examples

For example, case reports on complications of oral and peri-oral piercings (Hennequin-Hoenderdos et al. 2011), oral presentation of disseminated histoplasmosis (Akin et al. 2011), generalized severe gingival enlargement during pregnancy (McIntosh et al. 2010), periodontitis associated with leukocyte adhesion deficiency type I (Toomarian and Hashemi 2010), pemphigus vulgaris (Pradeep et al. 2010), or Chediak–Higashi syndrome (Khocht et al. 2010) can be found.

3.3.2 Case-Series Report

A case series aggregates individual cases in one report. Sometimes, the appearance of several similar cases in a short period heralds an epidemic (Grimes and Schulz 2002b).

Usually a case series involves 3–30 patients, all with a defined condition. To be worthy of publication, that condition must be either newly recognized, rarely seen, or highly educational. Case series are more compelling than a single case report because they demonstrate that the condition or some new treatment has occurred more than once. Case series can be strengthened by making them a comprehensive and consecutive collection of patients that includes all relevant cases within a set period and a given institution. Such steps limit concerns about selection bias. Case series do not have a concurrent control group; therefore, making comparisons or measuring associations is not possible (Thompson and Panacek 2007). A convenient feature of case-series reports is that they can constitute the case group for a case–control study, which can then explore hunches about causes of disease (Grimes and Schulz 2002b).

The strength of case reports and series lies in their unique ability to serve as an early warning system. New and unique diseases and their presentations are first observed here. Successful or failed treatments and complications are discovered. These studies are usually inexpensive and quickly performed. Of course, case reports and series do not reflect the general population. Frequently, they focus on exceptions to the usual presentation. Finally, data collection may not be adequate or well controlled (Luepker 2005).

There are papers on reporting of case series most notably by Moses (1984) and Abel (1998, 1999) and Jenicek's (2001) excellent recent book (Albrecht et al. 2005). They suggest, in essence, the following guidelines that need to be considered when writing and conducting case series (Albrecht et al. 2005):

1. Case series should be longitudinal, according to a predefined protocol. The protocol should define eligible patients, dosage, and treatment regime as precisely as possible. This reflects the thought process that should precede any treatment innovation.

2. It must explain why they have been observed, thus outlining basic inclusion and exclusion criteria. The reader has to understand why patients have been chosen. Patients selected due to failure of previous treatments and untreated volunteers have a different chance of cure; thus, the protocol must allow for these differences, and analysis must be separate, if necessary.
3. All patients in a department/an institution fulfilling the criteria and consenting should be treated according to plan.
4. The diagnosis of the patients must be well documented and open to scrutiny.
5. All patients should be observed with respect to outcome, which needs to be clearly defined and measured as objectively as possible, even those that may refuse the innovative treatment. An intention to treat analysis should be considered standard for non-randomized studies.
6. There should be some indication of treatment success rate in untreated or differently treated patients.
7. The place of treatment should be described, indicating differences of patient collectives, for example, between primary and tertiary centres.

Case series that are carefully designed and, of course, very similar to clinical trials are likely to minimize bias and maximize the information that can be deduced from a limited numbers of patients. Case reports are essentially the beginning of what, through duplication elsewhere, or in by the same physician, can become a case series and should therefore follow these rules where possible (Albrecht et al. 2005).

Examples

In the recent periodontal literature, a case series on periodontal status of patients with hypophosphatemic rickets (Ye et al. 2011), treatment of intrabony periodontal or furcation defects by bone grafting, guided tissue regeneration therapy, or enamel matrix derivative can be found (Okuda et al. 2009; Pretzl et al. 2008; Stavropoulos and Karring 2005; Donos et al. 2003; Zahedi et al. 2003; Rosen and Reynolds 2002; Kim et al. 2002).

Stavropoulos and Karring (2005) reported on the clinical and radiographic results 5 years following treatment of intrabony defects with guided tissue regeneration (GTR) in combina-

tion with deproteinized bovine bone (DBB) (Bio-Oss). Fifteen patients, with at least one intrabony periodontal defect with probing pocket depth (PPD)³7 mm and radiographic presence of an intrabony component (IC)³4 mm, were treated with a PLA/PGA bioabsorbable membrane (Fig. 3.5). The clinical measurements of each individual patient at baseline and at the 1- and 5-year control are presented in Table 3.3.

3.3.3 Surveys

The term survey describes a type of study that consists of asking people to respond to questions. Possible formats include personal interviews, telephone interviews, mailed questionnaires, Internet-based surveys, etc. (Totten et al. 1999). A more formal definition is ‘the ongoing systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know’ (Grimes and Schulz 2002b).

As a research tool, surveys have a number of important advantages. First, they are relatively inexpensive to perform. Second, they often allow quick data acquisition. Third, an appropriately drawn sample from a target population can provide accurate and representative data. Accurate results are certainly the goal of any research project, including survey-based research. There are some basic principles that apply to surveys that increase the likelihood of obtaining accurate and meaningful data. They are listed in Table 3.4 as the ‘ten commandments of performing survey-based research’. Poor quality survey research, performed by novice investigators, generally fails to apply these ten principles. Because of that, they often end up with results that are not valid (Totten et al. 1999).

Questions can take many forms, which have advantages and disadvantages. For example, questions can be open-ended or closed. They can have discrete or scalar or open responses. Open-ended questions elicit a ‘free text’ or essay-type response; closed questions must be answered by selected options. Open-ended questions encourage respondents to offer more information (which is why questions like ‘Why are you here today?’

are encouraged early in the medical interview), whereas closed questions simplify data input and analysis. Responses to open questions generally must be categorized by the researcher before analysis. Open questions are most useful during the initial study question design and development process. After the first few design iterations, the ability to use appropriate closed questions usually becomes apparent. Whenever possible, closed questions are preferred because they simplify result analysis (Totten et al. 1999).

A primary problem with surveys is **response bias**. In many cases, the individuals who return a survey vary greatly from those who do not; thus, the researcher cannot state with certainty that the results obtained represent the beliefs of the entire sample. The lower the rate of survey return encountered, the higher the chance that the results are biased. Because only the most ‘radicalized’ segment of the survey population may have responded to the survey, ask yourself the question, ‘If everyone who did not respond to the survey had responded with answers that were the opposite of the study findings, would it substantially change the study conclusions?’ If the answer to that is yes, then the results of the survey are at least suspect, if not invalid. A target response rate of 85% is a desirable goal because even if the remaining 15% had entirely different responses, they would be unlikely to change the overall conclusions substantially. However, response rates in most studies are generally much lower than 85%. Thus, increasing response rates is perhaps the most important method for increasing the validity of a survey research study (Thompson and Panacek 2007).

There are established techniques for improving response rates. Not all are applicable to every project. Consider the perspective of a respondent who receives the survey. What would make it more likely for them to complete it? (Panacek 2008b)

1. The *first* principle is to keep the survey short and focused. Many people will not answer a long survey, so shorten wherever possible. Focus on the questions truly necessary to answer the primary research question. Eliminate other peripheral or redundant questions.
2. *Second*, make the survey user-friendly. Poorly organized surveys, which are difficult to follow, are not well received. Computer software

Fig. 3.5 (a) A deep pocket is probed on the distal side of 46 at baseline. (b) A deep defect with evidence of an intrabony component (IC) >4 mm is observed in the baseline x-ray. (c) The presence of a predominantly 2-wall defect, with an $IC \geq 4$ mm, is confirmed during surgery after removal of the granulation tissue. (d) Prior to the final placement and stabilization of the membrane, Bio-Oss impregnated with gentamicin sulphate 2 mg/ml is loosely packed into the defect without overfilling it. (e) The clinical result 1 year after treatment. (f) Resolution of the defect with RBL gain and slight crestal resorption is observed in the x-ray 1 year after treatment. (g) The clinical result 5 years after treatment. Note the slight improvement on the soft tissue conditions when compared with the 1-year result. (h) Radiographically stable conditions are observed 5 years after treatment (Stavropoulos and Karring 2005. Reprinted with permission from Springer Science + Business Media)

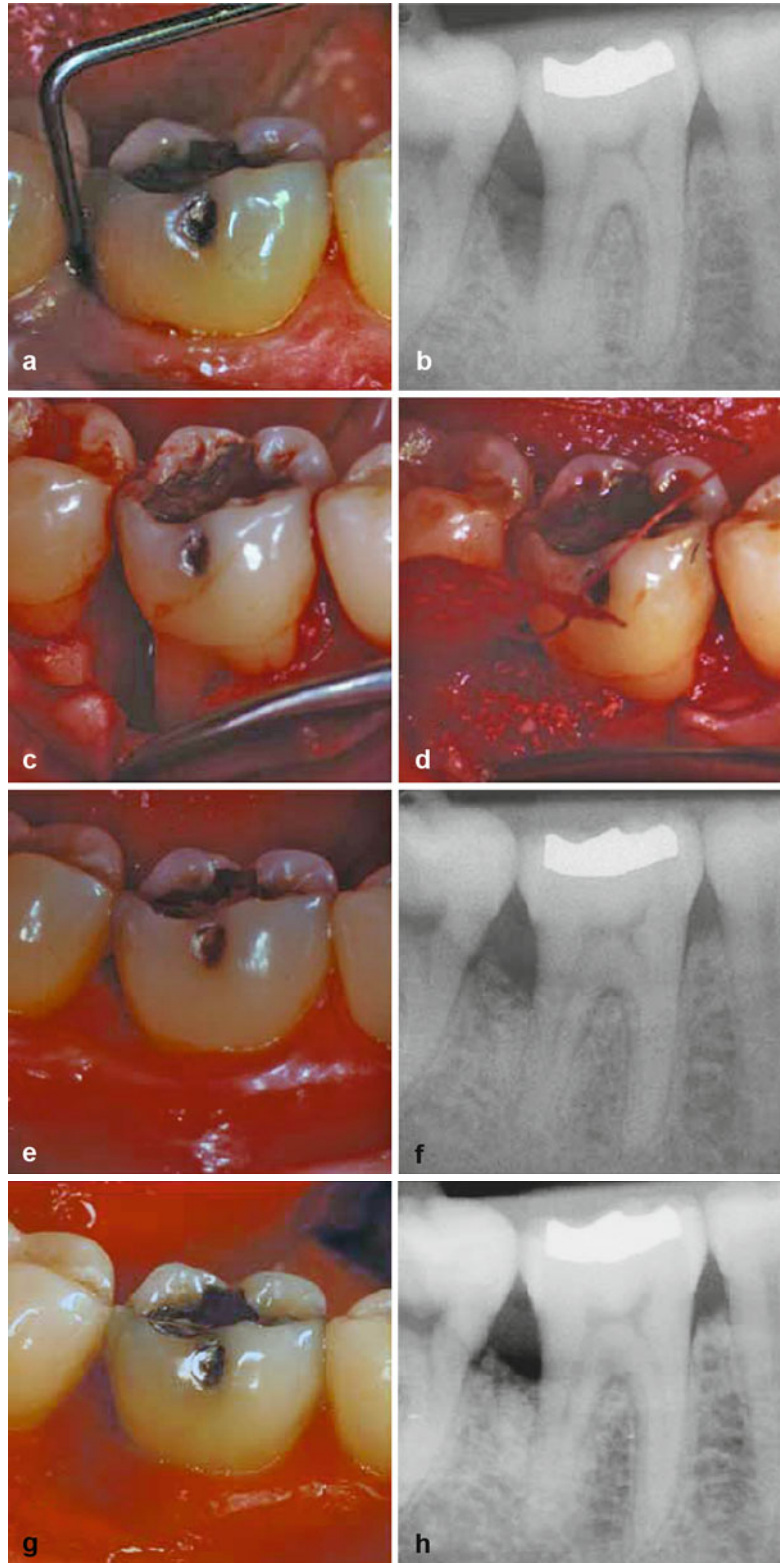


Table 3.3 Clinical measurements (in millimeter) of each individual patient at baseline and at the 1- and 5-year control (Stavropoulos and Karring 2005) (Reprinted with permission from Springer Science + Business Media)

Patient number	Baseline			1 year				5 years			
	PPD	REC	PAL	PPD	REC	PAL	PAL gain	PPD	REC	PAL	PAL gain
1. E.E.	9	0	9	3	0	3	6	4	0	4	5
2. P.N.	9	0	9	3	0	3	6				
3. A.A.	9	0	9	5	2	7	2	5	0	5	4
4. R.W.	8	1	9	2	6	8	1				
5. J.N.	7	1	8	4	2	6	2	3	3	6	2
6. H.K.	10	1	11	6	2	8	3				
7. L.V.	9	1	10	6	2	8	2				
8. U.H.	8	2	10	5	3	8	2	5	2	7	3
9. R.A.	9	0	9	3	3	6	3	4	1	5	4
10. I.W.	9	2	11	3	4	7	4	5	3	8	3
11. I.K.	9	0	9	4	1	5	4	5	0	5	4
12. H.H.	11	1	12	6	0	6	6	7	3	10	2
13. S.S.	11	0	11	5	1	6	5	6	0	6	5
14. P.J.	10	3	13	4	5	9	4	3	3	6	7
15. G.L.	10	2	12	4	1	5	7	4	2	6	6

PPD Probing pocket depth, REC gingival recession, PAL probing attachment level

programs make it easy to improve the design and readability of surveys. Take the time for that and solicit suggestions from volunteer reviewers.

3. *Third*, people respond better when they believe the survey is important. If you can generate a sense of importance for your study, response rates will increase.
4. *Fourth*, incentives can also help increase survey response rates. Small monetary gifts or gift certificates can be remarkably effective. However, this requires a study budget sufficient to support that option, which many novice researchers may not have.
5. *Fifth*, every survey that uses a mailing (direct postal or email) distribution should plan on second or even third mailings to increase the response rate. This should be built into the study plan. For example, if postage is involved, the study budget should include the cost of possible additional mailings.

Another important principle is, whenever possible, **to use existing validated instruments** rather than developing a new measurement scale in your survey. No matter how well you design a survey or questionnaire, potential misunderstandings always occur. When available, adapt prior surveys that have already dealt with this. Researchers should

Table 3.4 Ten commandments of surveys research (Panacek 2008a) (Reprinted with permission from Elsevier)

1. Start with a clear-cut research question
2. Keep the survey focused and user friendly
3. Use prior research methodology and questionnaires whenever possible
4. Decide on open versus closed question formats
5. Validate the survey instrument
6. Pilot test the survey instrument
7. Use proper sampling techniques
8. Plan for low response rates
9. Adjust for nonresponders
10. Do not overinterpret study results

always pilot test any new questions or surveys in a group of people resembling the desired sample. Revise the survey with feedback from the pilot test. If a large number of changes were made, the survey should be pilot tested a second or even third time to eliminate as much potential for confusion as possible (Thompson and Panacek 2007).

The term *validity* is the scientific design term for accuracy. In other words, does the measurement tool accurately measure what is intended? All studies measure something. Any survey instrument is a measurement tool designed to collect data accurately and reliably. However, unlike mechanical instruments such as blood pressure monitors or a

laboratory chemistry machine, the validity of a survey instrument is not automatically assumed. It is the investigator's responsibility to establish the validity of the survey instrument being used in their respective study (Panacek 2008a).

There are a number of different approaches to establishing validity. However, the easiest is to use or adapt an already validated survey. As mentioned, performing a quality literature search on the subject is important for multiple reasons. It provides guidance and insight into performing survey-based research and study methodology. Just as important, it may provide an actual survey instrument that can be directly used or modified for the purposes of the new investigation. Once a survey questionnaire has already been used in a prior published study, it is considered to be at least generally validated. If indeed it appeared to measure its intended target, it is accepted to be valid from a peer review standpoint. Therefore, use previously published survey instruments whenever possible. If they require some modification, attempt to keep those changes at a minimum and be prepared to argue that those changes did not substantially alter the survey instrument's overall validity (Panacek 2008a).

Often, there are no previously published survey instruments or questionnaires that are directly applicable to the investigator's intended project. It then is the responsibility of the investigator to validate the new survey instrument. Five general categories of validation are considered in such situations (Table 3.5) (Panacek 2008a).

Examples

Deinzer et al. (2009) assessed periodontitis-related knowledge and its relation to oral health behaviour on a community level and to identify target groups and major topics for health education interventions. By means of a multistratified, stochastic telephone survey, 1,001 interviews with Germans older than 14 years were carried out. The sampling model was based on all registered fixed network telephone numbers in Germany in 2006; an additional procedure allowed contact with non-registered telephone numbers as well. The survey was based on the Infratest Telephone Master Sample, a multistratified sample with area

Table 3.5 Categories of validation for survey instruments (Panacek 2008a) (Reprinted with permission from Elsevier)

Face validity	Does it make sense simply by looking at it?
Content validity	A panel of experts reviews and revises it for relevant content.
Criterion validity	How does it compare with an established standard in the field?
Predictive validity	How does it compare to actual outcomes of interest?
Construct validity	Does the internal construct/structure conform to scientific survey study principles?

sampling (geographic location of the population units in Germany) as a first step, a consecutive (second step) random digit dialling, and finally a stochastic selection of the person to be interviewed within a private home. Participants answered questions on the definition, aetiology, and risk factors of periodontal disease and on the risks associated with and measures to prevent them. Severe knowledge deficits were found with respect to all topics (Table 3.6). No consistent relationships with age or education were found, although less educated and very young and old people tended to show the greatest deficits. Knowledge of preventive measures was most strongly related to current oral health behaviour (Table 3.7).

Yuen et al. (2009) determined levels of oral health knowledge and factors associated with adequate oral health knowledge in adults with diabetes. A convenience sample of 253 adult US residents with diabetes completed an oral health survey to assess their knowledge. The questionnaire contained a combination of open- and close-ended questions (multiple choice and true/false format). In addition to the socio-demographic information, participants were asked to answer questions related to source of oral health information, oral health knowledge, oral health behaviour, dentition status, types of diabetes, years since diabetes diagnosis, and whether they had attended a diabetes education session in the past. Socio-demographic characteristics considered were age, gender, race, marital status, education, employment status, presence or absence of dental insurance, and annual household income. Oral

Table 3.6 Statements related to preventive measures of periodontitis: overall percentage of male and female participants deciding correctly whether the statement was right or wrong (Deinzer et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Statement	Men (%)	Women (%)	All (%)	Gender difference	
				χ^2 (df=1)	P
Statements related to oral hygiene					
• Periodontal disease is caused by dental plaque (right)	76.4	79.1	77.8	1.063	0.303
• If one manages to sustain very good oral hygiene he or she will not suffer from periodontitis (right)	71.2	64.4	67.6	5.311	0.021
• One cannot avoid emergence of dental calculus (wrong)	40.3	42.7	41.6	0.606	0.436
• To really get the teeth clean by daily brushing above all one has to brush them firmly (wrong)	62.4	70.7	66.7	7.692	0.006
• The fewest patients manage to sustain optimal oral hygiene without the help of their dentist (right)	74.3	67.2	70.6	6.036	0.014
• To avoid periodontitis, it is of particular importance to brush the chewing surfaces (wrong)	38.4	39.8	39.2	0.215	0.643
Statements related to the early diagnosis of periodontitis					
• The most frequent oral disease in adults is caries (wrong)	24.6	33.7	29.4	9.932	0.002
• In early stages one recognizes periodontitis by frequent tooth aches (wrong)	43.2	37.9	40.5	2.895	0.089
• A periodontitis often remains unrecognized for years (right)	72.0	71.8	71.9	0.004	0.948

Table 3.7 Differences in knowledge between daily users ($\geq 1/\text{day}$) of interdental devices and non-daily or never users: percentage of participants giving the correct answer (only differences with a $P \leq 0.05$ are reported) (Deinzer et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Statement	Non-daily/never users (n = 613) (%)	Daily users (n = 387) (%)	Group difference	
			χ^2 (df = 1)	P
• Active knowledge on hygiene as risk factor	27.7	38.2	12.202	<0.001
Statements related to oral hygiene				
• To really get the teeth clean by daily brushing above all one has to brush them firmly (wrong)	60.3	77.0	29.971	<0.001
• To avoid periodontitis it is of particular importance to brush the chewing surfaces (wrong)	35.7	44.7	8.133	0.004
Statements related to the early diagnosis of periodontitis				
• The most frequent oral disease in adults is caries (wrong)	24.3	37.5	19.940	<0.001
• In early stages one recognizes periodontitis by frequent tooth aches (wrong)	55.7	65.6	9.722	0.002
• Active knowledge on necessity of daily ($\geq 1/\text{day}$) interdental hygiene	54.5	97.9	218.833	<0.001

health behaviours included frequency of dental visits, brushing, flossing, and mouth rinse use. Administration of the questionnaire was conducted either individually or in small groups. Items and response choices were read to those participants who expressed having difficulty reading (i.e. illiteracy or poor vision). Results showed that only 47% of the participants answered five or more (out of a maximum of seven) oral health knowledge items related to diabetes correctly. Participants who received oral health information related to diabetes have 2.9 times the odds of possessing adequate oral health knowledge (i.e. answered five or more items correctly) compared to participants who did not received that information controlling for education and race (OR=2.86, 95% CI 1.31–6.24, $P=0.008$).

Schröen et al. (2011) evaluated the range of application, frequency of utilization, and subjective outcomes in the use of EMD during routine clinical practice among the Swiss specialists in periodontology (SP) and to compare their professional attitude to the non-specialist (NSP) but active members of the Swiss Society of Periodontology (SSP). A cross-sectional postal survey of 533 dentists, representing all members of the SSP practising in Switzerland, was conducted. All addressees were asked to respond within 3 months. A reminder telephone call was made after 2 months. All questionnaires, returned within 3 months, were included in the analysis. The questionnaire consisted of three sections, assessing (1) general personal information regarding the practice setting and education, (2) general questions regarding periodontal surgery practices, and (3) specific questions regarding the use of EMD. The information obtained was compared, and differences between specialists and non-specialists were calculated. Sixty-nine percent of the specialists answered the questionnaire, compared to only 37.4% of the non-specialists (overall 42.4%). In general, specialists performed surgeries more frequently and presented a significantly higher percentage of EMD users than the non-specialists. The survey revealed that EMD was used on a regular basis by dentists performing periodontal therapy. In addition, the answers by both groups generally corresponded well with the current available literature.

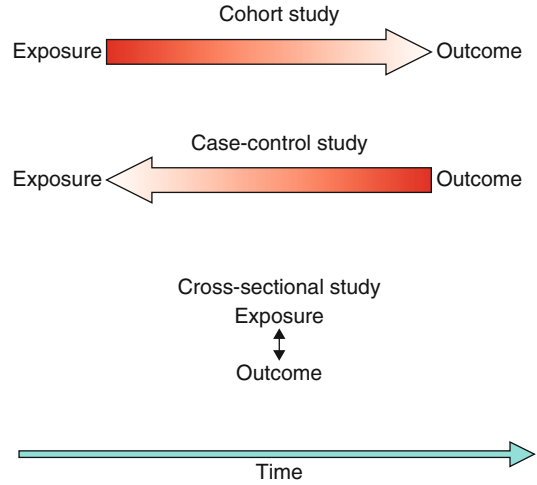


Fig. 3.6 Schematic diagram showing temporal direction of three study designs (Grimes and Schulz 2002a), (Reprinted from *The Lancet* 2002;359(9300):57–61 with permission from Elsevier)

3.3.4 Cross-Sectional Studies

Sometimes termed a frequency survey or a prevalence study, cross-sectional studies are done to examine the presence or absence of disease and the presence or absence of an exposure at a particular time. Thus, prevalence, not incidence, is the focus. Since both outcome and exposure are ascertained at the same time (**Fig. 3.6**), the temporal relation between the two might be unclear (Grimes and Schulz 2002a).

How to run a cross-sectional study? Formulate the research question(s) and choose the sample population. Then decide what variables of the study population are relevant to the research question. A method for contacting sample subjects must be devised and then implemented. In this way, the data are collected and can then be analysed (Mann 2003).

Such studies may be analysed like a cohort study by comparing disease prevalence between exposure groups. They may also be analysed like a case-control study by comparing the odds of exposure between groups with and without disease. A difficulty that can occur in any design but is particularly clear in cross-sectional studies is to establish that an exposure preceded the disease,

although the time order of exposure and outcome may sometimes be clear. In a study in which the exposure variable is congenital or genetic, for example, we can be confident that the exposure preceded the disease, even if we are measuring both at the same time (Vandenbroucke et al. 2007).

The **strengths** of a cross-sectional study are found in several outcomes. They are useful in determining the burden of disease in a population as a means of informing health-care services. They are useful in the generation of hypotheses for exposure/disease relationships. They are most useful in studying diseases of long duration, and they are relatively inexpensive (Luepker 2005).

Examples

Several cross-sectional studies that evaluated the association between periodontitis and coronary heart disease (CHD) (Mattila et al. 1989a, b; Pearson et al. 2003). Table 3.8 depicts the observational studies investigating the association between overweight or obesity (as defined by the World Health Organization) and periodontitis; various measures of periodontal disease can be identified (Suvan et al. 2011).

3.3.5 Case-Control Studies

A case-control design selects two similar populations of patients based on their outcome. One has the dependent variable of interest (i.e. the outcome), and the other does not. The investigator then looks backward, retrospectively, for the independent variables of interest, that is, looking for the causes for that outcome. This is the opposite of a cohort design in which you identify patient populations based on the independent variables being present (or not) and then generally look prospectively for the subsequent outcomes (Thompson and Panacek 2007). If the prevalence of the exposure is higher among cases than among controls, then the exposure is associated with an increased risk of the outcome (Grimes and Schulz 2002a).

Case-control designs might seem easy to understand, but many clinicians stumble over them. Because this type of study runs backwards by comparison with most other studies, it often confuses researchers and readers alike. In cohort studies, for example, study groups are defined by exposure. In case-control studies, however, study groups are defined by outcome (Figs. 3.7, 3.8, and 3.9) (Schulz and Grimes 2002).

Confusion about the term case-control study was found in medical literature. In Hellems's et al. (2006) analysis of 91 self-defined case-control studies published in two paediatric journals over an 8-year period, 68 of the identified articles met the standard definition of a case-control study. The other 23 articles could be classified as cross-sectional studies or cohort studies, although each was literally a case-control study because it contained both cases and controls. Mihailovic et al. (2005) found that only 35% of self-identified case-control studies published in six general surgical journals actually met the standard definition for this study design. The proportion of mislabelled case-control studies in rehabilitation journals was 97% (83 of 86 studies were mislabelled) (Mayo and Goldberg 2009). In the 124 reports identified in four US obstetrics and gynaecology journals, the proportion of mislabelled 'case-control' studies was 30% overall (Grimes 2009).

Particular types of research questions may lend themselves most aptly to particular study designs. Case-control studies, for example, can be efficient for assessing **outcomes that are rare** or that **take a long time to develop**. Since case-control studies do not require the assembly and prolonged follow-up of a large cohort of patients, they are typically much less expensive than cohort studies. The case-control design also allows one to study the effect of exposures that might be ethically untenable in an interventional trial, for example, the risk of exposure to tobacco smoke. Case-control studies are also well suited to exploring multiple etiologic hypotheses (i.e. multiple antecedent risk factors) since once the cases and controls are selected, it is efficient to inquire (or consult records) about multiple past exposures (Hellems et al. 2006).

Table 3.8 Characteristics of included cross-sectional studies included in a systematic review investigating the association between overweight/obesity and periodontitis in adults. (Suvan et al. 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Author (country)	Population/sample	Body composition (exposure)	Periodontal outcome (dependent variables)	Covariates included for adjustment	Publication conclusions	Effect OR (95% CI) or other as stated
Haffajee and Socransky (2005) (USA)	Retrospective analysis of subjects included in previous trials at The Forsyth Institute $n=695$ Periodontally healthy/gingivitis = 121 Chronic periodontitis = 574 18–86 years	BMI WHO categories Underweight < 18.5 kg m ⁻² Normal 18.5–24.9 (reference) Overweight 25–29.9 Obese ≥ 30 Body fat %	Full mouth assessment of plaque, bleeding, PD, and CAL Periodontally healthy/gingivitis defined as minimum 24 teeth present with <2% of sites with PD ≥ 4 mm Chronic periodontitis defined as minimum 20 teeth present with ≥ 5% of sites with PD ≥ 4 mm and/or >5% sites with CAL ≥ 4 mm	Age Gender	Prevalence Subjects, particularly younger females who were overweight or obese, were at greater risk for periodontitis than subjects with normal BMI after adjusting for age, gender and smoking	BMI and risk for periodontitis (ref. normal) Overweight 3.07 (1.91, 4.84) Obese 5.31 (1.41, 1.644) BMI adjusted for gender and age (ref. overweight and normal) Obese 2.31 (1.19, 4.49) $P=0.0086$ Periodontally healthy/gingivitis Normal 65.3% Overweight 24.0% Obese 10.7% Chronic periodontitis Normal 32.9% Overweight 37.8% Obese 29.3% Mean BF % Healthy 25.0 ± 10.2 Periodontitis 29.5 ± 10.2
Han et al. (2010) (Korea)	Siwha-Banwol Environmental Health Cohort $n=1046$ 15–84 years (separate data reported on adults 45–54 years)	BMI Overweight 23–25 Obese ≥ 25 Waist Circumference ≥ 85 high for women ≥ 90 high for men WHR ≥ 0.85 high for women ≥ 0.90 high for men VFA Obesity ≥ 100 cm ²	CPITN PPD scored one code per sextant based on deepest pocket using WHO probe CPITN Code 0–2 no periodontitis (PPD ≤ 3.5 mm) CPITN Code 3 or 4 periodontitis (PPD ≥ 4.0 mm)	Age, gender, monthly family income, smoking, frequency of daily brushing, physical activity	Prevalence Obesity was associated with periodontitis	BMI and risk for periodontitis (ref. normal) Males 45–54 years ($n=96$) Overweight 1.11 (0.33, 3.74) Obese 2.50 (0.85, 7.38) Females ($n=102$) Overweight 1.14 (0.39, 3.28) Obese 1.47 (0.52, 4.16) WC Males 45–54 years ($n=96$) Obese 1.46 (0.60, 3.59) Females ($n=102$) Obese 2.76 (1.05, 7.27)
Hilgert et al. (2009) (Brazil)	$n=783$ Elderly Southern Brazilian people 60–91 years	BMI WHO category obese Non-obese < 30 Obese ≥ 30	Examined for number of teeth and use of prostheses	Demographic and socio-economic parameters and smoking	Poorer oral status represented by tooth loss was associated with obesity in Southern Brazil older people	BMI and risk for upper denture only OR (ref. non-obese) Obese 2.34 (1.18, 4.27) BMI and risk of ≤ 8 teeth present OR (ref. non-obese) Obese 2.96 (1.68, 5.19)
Khader et al. (2009) (Jordan)	Individuals accompanying patients during their visit to outpatient clinics to Jordan University Medical Centre $n=340$ 18–70 years	BMI WHO categories Underweight < 18.5 kg m ⁻² Normal 18.5–24.9 (reference) Overweight 25–29.9 Obese ≥ 30 WC > 88 high for women > 102 high for men BF % < 30% normal > 30% high WHR obesity > 0.90 men > 0.87 women	Full mouth assessment for PPD, CAL, PII and GI on Ramfjord teeth Periodontitis defined as presence of > 4 teeth with minimum one site with PPD ≥ 4 mm and CAL ≥ 3 mm	Age, plaque index, number of missing teeth	Prevalence BMI defined obesity, high WC, and high fat per cent were significantly associated with increased odds of having periodontitis	BMI and risk for periodontitis (ref. normal) Overweight 1.40 (0.70, 3.0) $P=0.330$ Obese 2.9 (1.3, 6.1) $P=0.006$ WC and risk for periodontitis (ref. normal) High 2.1 (1.2, 3.7) $P=0.009$ WHR and risk for periodontitis (ref. normal) High 1.4 (0.8, 2.4) $P=0.221$ Body Fat% and risk for periodontitis (ref. normal) High 1.8 (1.03, 3.3) $P=0.039$ Presence of periodontitis in BMI categories Normal 14.0% Overweight 29.6% Obese 51.9% $P<0.005$ Presence of periodontitis in WC categories Normal 19.5% High 46.5% $P<0.005$ Presence of periodontitis in WHR categories Normal 24.7% High 38.2% $P=0.008$

Kongstad et al. (2009) (Denmark)	Subset of Copenhagen Heart Study 2004–2007 <i>n</i> = 1,504 20–95 years	BMI – self report WHO categories Underweight < 18.5 kg m ⁻² Normal 18.5–24.9 (reference) Overweight 25–29.9 Obese ≥ 30	Full mouth assessment for PPD, CAL, BOP, and PI Full mouth mean CAL dichotomized as ≥ 3 mm < 3 mm	Age, gender Or Age, gender, smoking, educational level, income, alcohol consumption, physical activity, diabetes, number of teeth, and plaque score	Prevalence BMI be inversely associated with clinical AL but positively related to BOP	BMI and risk for periodontitis (ref. normal) CAL ≥ 3 mm adjusted for age and gender Overweight 1.03 (0.76, 1.39) Obese 1.08 (0.71, 1.64) CAL ≥ 3 mm adjusted for all covariates Overweight 0.87 (0.61, 1.23) Obese 0.60 (0.36, 0.99)
Kushiyama et al. (2009) (Japan)	<i>n</i> = 1,070 residents of Miyazaki City (subset of Health Study who had oral examination) 281 males, 789 females 40–70 years	BMI ≥ 25 Overweight and Obese < 25 Non-overweight or obese	CPITN PPD scored one code per sextant based on deepest pocket using WHO probe CPITN Code 0–2 no periodontitis (PPD ≤ 3.5 mm) CPITN Code 3 or 4 periodontitis (PPD ≥ 4.0 mm)	Age, gender and smoking	Prevalence Study results support suspected relationship between metabolic syndrome and periodontal disease	BMI and risk for periodontitis CPITN = 4 (BMI ref. < 25) Overweight and Obese 1.09 (0.77, 1.53) <i>P</i> = 0.64
Nucci Da Silva et al. (2009) (Brazil)	<i>n</i> = 214 29–65 years	BMI WHO categories Underweight < 18.5 kg m ⁻² Normal 18.5–24.9 (reference) Overweight 25–29.9 Obese ≥ 30 WC WHR BF %	Full mouth periodontal examination Periodontitis defined as presence ≥ 4 teeth with minimum one or more sites with probing PD ≥ 4 mm and CAL ≥ 3 mm	Not reported	Prevalence BMI defined obesity, high WC, and high fat percent were significantly associated with increased odds of having periodontitis	Prevalence of periodontitis by BMI category Normal 9.5% Overweight 39.6% Obese 51.9% BMI and risk for periodontitis OR (ref. normal) Obese 2.7 (1.8, 6.4)
Östberg et al. (2009) (Sweden)	<i>n</i> = 2761 Data from Skaraborg Project 2001–2005 30–74 years	BMI WHO categories Underweight < 18.5 kg m ⁻² Normal 18.5–24.9 (reference) Overweight 25–29.9 Obese ≥ 30 WC > 88 high for women > 102 high for men WHR BF %	Tooth loss dichotomized as 0–19 teeth remaining or > 20 teeth remaining (evaluated by patient questionnaire)	Age, gender	Tooth loss is associated with both general and abdominal obesity in participants < 60 years	BMI risk of < 20 teeth remaining (OR) Obese 1.25 (0.94, 1.65) WC risk of < 20 teeth remaining (OR) Obese 1.43 (1.10, 1.85)
Saxlin et al. (2009) (Finland)	Finland Health 2000 Survey <i>n</i> = 425 45–64 years Never smoker	BMI Used BMI as continuous measure	Full mouth assessment for PPD, CAL Analysed by number of pockets > 4 and > 6 mm	Age, gender, education, presence of plaque and number of teeth	Prevalence Serum IL-6 was associated with periodontal infection	BMI and no. teeth with PPD ≥ 4 mm RR = 1.03 (1.00, 1.06)

(continued)

Table 3.8 (continued)

Author (country)	Population/sample	Body composition (exposure)	Periodontal outcome (dependent variables)	Covariates included for adjustment	Publication conclusions	Effect OR (95% CI) or other as stated
Ekuni et al. (2008) (Japan)	Subset of Health and Medical Center of Okayama University health survey who had dental examination $n=618$ 18–24 years Males = 296 Females = 322	BMI WHO categories: Underweight < 18.5 kg m ⁻² Normal 18.5–24.9 (reference) Overweight 25–29.9 Obese ≥ 30 BF %	CPITN PPD scored one code per sextant based on deepest pocket using WHO probe CPITN Code 0–2 no periodontitis (PPD ≤ 3.5 mm) CPITN Code 3 or 4 periodontitis (PPD ≥ 4.0 mm)	Age, gender	Prevalence BMI could be a potential risk factor for periodontitis among healthy young individuals	Risk for periodontitis per 1 kg m ⁻² 16% increased risk for periodontitis per 1 kg m ⁻² Adjusted OR 1.16 (1.03, 1.31)
Ylösto et al. (2008) (Finland)	Dentate, non-diabetic subset of Health 2000 Health Examination Survey $n=2,842$ 30–49 years	BMI Categorized in quintiles: <22.3 22.3–24.3 24.3–26.3 26.3–29.1 >29.1	WHO probe measuring four sites per tooth recording only deepest site per tooth Analysed as number of teeth with probing pockets of 4 mm or deeper and 6 mm or deeper	Age, gender, smoking, education, dental attendance pattern, tooth brushing frequency, presence of plaque and no. of teeth	Prevalence, Extent and Severity Results showed an association between body weight and periodontal infection among the non-diabetic, non-smoking population aged 30–49	RR for BMI and no. teeth with PD > 4 mm in total population: <22.3 reference 22.3–24.3 1.1 (0.9, 1.3) 24.3–26.3 1.0 (0.8, 1.1) 26.3–29.1 1.1 (0.9, 1.2) >29.1 1.2 (1.0, 1.4) Continuous variable 1.01 (1.00, 1.02) RR for BMI and no. teeth with PD > 6 mm in total population: <22.3 reference 22.3–24.3 0.9 (0.6, 1.4) 24.3–26.3 0.6 (0.4, 1.0) 26.3–29.1 0.8 (0.5, 1.3) >29.1 1.1 (0.7, 1.6) Continuous variable 1.01 (0.96, 1.05)
Shimazaki et al. (2007) (Japan)	Subset Hisayama community health study $n=584$ women with minimum 10 teeth 40–79 years	WC Categorized as ≤88 reference >88 high	PD and clinical attachment level measured mesio-buccal and mid-buccal site of all teeth in two quadrants (one max, one mand) Categorized using mean as PD <2.0 or ≥2.0 mm and CAL <3 mm or ≥3.0 mm	Age and smoking status	Prevalence Results indicated an association between components of metabolic syndrome and periodontitis in a female population	Risk for mean PD ≥ 2.0 mm with WC > 88 (OR) 1.8 (1.2, 2.8) Risk for mean CAL ≥ 3.0 mm 0.9 (0.4, 1.9) Mean (SD) WC in groups by: Mean PD < 2.0 WC 82.7 (10.1) Mean PD ≥ 2.0 WC 85.8 (8.9) $P=0.01$ Mean CAL < 3.0 WC 83.2 (10.0) Mean CAL ≥ 3.0 WC 84.6 (9.8) NS
Borges-Yanez et al. (2006) (Mexico)	Older age subset of Cross-cultural research on nutrition study $n=473$ non-diabetic >60 years	BMI Obesity defined as BMI ≥ 30	PD and clinical attachment level measured mesio-buccal and mid-buccal site of all teeth in two quadrants (one max, one mand) Moderate periodontitis – minimum two sites with attachment loss ≥ 4 mm Severe periodontitis – minimum one site with attachment loss ≥ 6 mm	Calculus index	Prevalence Combined obesity and high calculus index associated with high probability of having periodontitis	BMI ≥ 30 Risk (OR) for moderate or severe periodontitis 0.9325 (0.8524, 1.0170) BMI and high calculus index 1.1367 (1.0000, 1.3058) No periodontitis Mean BMI 25.6 (SD 4.84) $n=152$ Moderate or Severe Periodontitis Mean BMI 25.7 (SD 4.9) $P=0.084$ $n=169$

Al-Zahrani et al. (2005) (USA)	NHANES III subset with periodontal measurements and similar physical activity level for ≥ 10 years Non-diabetic $n = 2,521$ >18 years	BMI	PD and clinical attachment level measured mesio-buccal and mid-buccal site of all teeth in two quadrants (one max, one mand) Periodontitis classed as minimum one site with attachment loss ≥ 3 mm and PPD ≥ 4 mm	Age, gender, smoking, race, poverty index, education, vitamin use, healthy eating index, time since last dental visit, gingival bleeding, dental calculus	None stated relating to body composition Prevalence results reported	Bivariate association of BMI and periodontitis No periodontitis Mean BMI 25.8 (se 0.2) Periodontitis Mean BMI 26.6 (se 0.3) $P = 0.039$ Multivariable association. OR 1.01 (0.98, 1.05) P value 0.524
Chapper et al. (2005) (Brazil)	Sample from Gestational Diabetes Outpatient Unit $n = 60$ women with gestational diabetes Non smokers 19–45 years $n = 18$ normal BMI $n = 15$ Overweight $n = 27$ Obese	BMI WHO categories 18.5–24.9 normal 25–29.9 overweight ≥ 30 obese	CAL recording deepest site of six/tooth GB and BOP	Only female, non smokers with diabetes in sample	Severity Patients with gestational diabetes and pregestational obesity had more gingival inflammation and attachment loss than those with normal pregestational BMI	GB (mean, SD) Normal 52.76 $\pm$ 27.99* Overweight 65.64 $\pm$ 23.31 Obese 78.85 $\pm$ 27.44* $P < 0.05$ BOP (mean, SD) Normal 55.65 $\pm$ 27.65 Overweight 71.47 $\pm$ 20.35 Obese 75.31 $\pm$ 30.33 $P = 0.062$ CAL (mean, SD) Normal 2.21 $\pm$ 0.41* Overweight 2.40 $\pm$ 0.53 Obese 2.61 $\pm$ 0.54* $P < 0.05$ No. of teeth (mean, SD) Normal 24.44 $\pm$ 6.43 Overweight 25.27 $\pm$ 5.24 Obese 26.25 $\pm$ 5.10 $P = 0.556$
Dalla Vecchia et al. (2005) (Brazil)	Subset of a larger representative sample of metropolitan area of Porto Alegre, RioGrande do Sul, Brazil $n = 706$ 30–65 years	BMI Categorized as: <18.5 underweight 18.5–24.9 normal 25–29.9 overweight ≥ 30 obese	Full mouth periodontal assessment (PPD and CAL) at six sites per tooth Presence of periodontitis classed as $\geq 20\%$ teeth with attachment loss ≥ 5 mm	Male and Female separate analysis Age, smoking, race, socioeconomic status, presence of calculus, frequency of dental visits	Prevalence Obesity was significantly associated with occurrence of periodontitis in adult, non-smoker women	BMI Risk for periodontitis Males ($n = 329$) Normal reference Overweight 1.1 (0.4, 3.3) Obese 1.0 (0.5, 1.7) Females ($n = 377$) Normal reference Overweight 1.3 (0.8, 2.2) Obese 2.1 (1.1, 3.9) $P < 0.05$
Genco et al. (2005) (USA)	Part I NHANES III $n = 12,367$ non-diabetic 20–90 years	Part I BMI ≥ 27 Overweight <27 Non-overweight	Part I PD and attachment level measured mesio-buccal and mid-buccal site of all teeth in two quadrants (one max, one mand) Mean attachment loss ≥ 1.5 mm defined as periodontal case	Age, gender, smoking, education, race/ethnicity	Prevalence Obesity is a significant predictor of periodontal disease	Part I Risk (OR) for periodontitis BMI ≥ 27 1.48 (1.13, 1.93) Mean % individuals with periodontal disease in groups: ≥ 27 Overweight 24.3 <27 Non-overweight 26.9

(continued)

Table 3.8 (continued)

Author (country)	Population/sample	Body composition (exposure)	Periodontal outcome (dependent variables)	Covariates included for adjustment	Publication conclusions	Effect OR (95% CI) or other as stated
Nishida et al. (2005) (Japan)	<i>n</i> = 372 factory workers with complete data from a sample of 409 20–59 years	BMI subdivided in six groups <20 20–21.9 22–23.9 24–25.9 26–27.9 ≥28	PPD assessed at six sites/teeth, only deepest site/teeth recorded Calculated % of teeth with PPD > 3.5 mm Periodontitis = % teeth with PPD > 3.5 mm ≥20% Non-periodontitis = % teeth with PPD > 3.5 mm <20%	Age, gender, smoking, alcohol consumption, frequency of tooth brushing	Prevalence Smoking displays the greatest impact on periodontitis among lifestyle-related factors. Both smoking and obesity are independent risk indicators for periodontitis; moreover, these parameters exhibit a dose-response relationship with respect to periodontitis risk	BMI Risk (OR) for periodontitis BMI ≥ 25 Reference BMI < 25 3.17 (1.79, 5.61) Risk for periodontitis by BMI category <20 reference 20–21.9 1.88 (0.66, 5.37) 22–23.9 1.74 (0.61, 4.97) 24–25.9 2.68 (0.90, 7.98) 26–27.9 3.89 (1.20, 12.56) ≥28 4.40 (1.18, 16.44) <i>P</i> value for trend <i>P</i> = 0.0057
Saito et al. (2005) (Japan)	Subset Hisayama community health study <i>n</i> = 584 women with minimum 10 teeth 40–79 years	All measures divided into quartiles for analysis BMI WHR BF (measured by bio-impedance)	PD and clinical attachment level measured mesio-buccal and mid-buccal site of all teeth in two quadrants (one max, one mand) Categorized in quintiles based on calculated mean PPD and CAL	Age, smoking, plaque index, socioeconomic status, glucose tolerance test	Prevalence Obesity was associated with deep pockets in Japanese women, even after adjusting for oral glucose tolerance test	Highest quintile PPD (≥1.9 mm) and BMI quartile BMI 20.8–22.7 3.0 (1.4, 6.3) <i>P</i> < 0.01 BMI 22.7–24.9 2.3 (1.1, 5.0) <i>P</i> < 0.05 BMI 25.0–46.7 4.3 (2.1, 8.9) <i>P</i> < 0.001 Highest quintile PPD (≥1.9 mm) and Body Fat quartile BF 24.2–27.9 2.6 (1.2, 5.3) <i>P</i> < 0.05 BF 28.0–32.5 2.8 (1.3, 5.7) <i>P</i> < 0.01 BF 32.6–52.5 3.3 (1.6, 6.8) <i>P</i> < 0.01 Highest quintile PPD (≥1.9 mm) and WHR quartile WHR 0.89–0.94 1.4 (0.7, 2.8) WHR 0.94–0.97 2.8 (0.6, 2.4) WHR 0.97–1.12 2.1 (1.1, 4.1) <i>P</i> < 0.05 Highest quintile CAL (≥2.42 mm) and BMI quartile BMI 20.8–22.7 1.6 (0.8, 3.1) BMI 22.7–24.9 1.3 (0.7, 2.6) BMI 25.0–46.7 1.8 (0.9, 3.4) Mean (SD) body composition in groups by: Mean PD < 1.9 BMI 22.9 (3.5) BF 28.0 (6.1) WHR 0.93 (0.06) Mean PD ≥ 1.9 BMI 24.1 (2.9) <i>P</i> = 0.0004 BF 30.4 (5.7) <i>P</i> = 0.0002 WHR 0.94 (0.05) <i>P</i> = 0.027

Torrungruang et al. (2005) (Thailand)	$n=2,005$ participants 50–73 years Survey of senior employees and retired personnel group of Thai workers	BMI Categorized as <18.5 underweight 18.5–22.9 normal 23–24.9 at risk 25–29.9 obese I ≥ 30 obese II WC With cut-off point of ≥ 80 for women and ≥ 90 for men	Two randomly selected quadrants Plaque index (PI) plaque present or absent, two buccal sites/tooth one quad and two lingual sites/tooth one quad PD six sites/tooth CAL six sites/tooth Categorized as Mild CAL <2.5 mm (reference) Moderate CAL 2.5–3.9 mm Severe CAL ≥ 4.0 mm	Age, gender, oral hygiene status, education level, income, smoking status, alcohol consumption, diabetes mellitus	Severity Age, gender, education, oral hygiene status, smoking, and diabetes are significantly associated with periodontal disease severity in this study group	BMI Risk for moderate periodontitis normal 1.1 (0.8, 1.4) obese I 1.0 (0.8, 1.3) obese II 1.4 (0.9, 2.3) High WC risk for moderate periodontitis (0.8–1.2) BMI Risk for severe periodontitis normal 1.0 (0.7, 1.4) obese I 1.0 (0.8, 1.5) obese II 1.5 (0.8, 2.7) High WC risk for moderate periodontitis 0.8 (0.6–1.1)
Al-Zahrani et al. (2003) (USA)	NHANES III subset with complete periodontal measurements $n=13,665$ 18–90 years	BMI Categorized as: <18.5 underweight 18.5–24.9 normal 25–29.9 overweight ≥ 30 obese WC >88 high for women >102 high for men	PD and clinical attachment level measured mesio-buccal and mid-buccal site of all teeth in two quadrants (one max, one mand) Periodontitis classed as minimum one site with attachment loss ≥ 3 mm and PPD ≥ 4 mm	Age, gender, smoking, race, poverty index, education, diabetes, time since last dental visit	Prevalence BMI > 30 and WC > 88 cm associated with increased prevalence of periodontal disease in younger population	BMI Overweight 1.06 (0.912, 1.235) $P=0.434$ Obese 1.37 (1.41, 1.644) $P=0.001$ WC High 1.33 (1.113, 1.601) $P=0.002$ BMI Stratified by Age 18–34 Overweight 1.00 (0.705, 1.407) Obese 1.76 (1.187, 2.612) $P<0.01$ 35–59 Overweight 0.98 (0.812, 1.191) Obese 1.12 (0.820, 1.522) 60–90 Overweight 1.20 (0.882, 1.626) Obese 1.20 (0.892, 1.608) Prevalence of Periodontal Disease % (SE) <18.5 underweight 9.15 (2.488) $n=301$ 18.5–24.9 normal 11.55 (0.828) $n=5,289$ 25–29.9 overweight 14.59 (0.856) $n=4,623$ ≥ 30 obese 17.65 (1.008) $n=3,427$
Wood et al. (2003) (USA)	Caucasian, >18 years, participants of NHANES III (sample of general population) $n=8,032$ overall (varied by availability of outcome data)	BMI WHR Subcutaneous fat (S) Fat-free mass (FFM)	PD and clinical attachment level recorded mesial and mesio-buccal in two contra-lateral quadrants	Age, gender, history of diabetes, current smoking, and socioeconomic status	Prevalence, and Severity Significant correlations between body composition and periodontal disease (with WHR most significant, followed by BMI, FFM, and S) showed similarities to those observed in other obesity-related health problems	Correlation coefficient: Between mean periodontal attachment loss and BMI 0.1512 ($P<0.01$) WHR 0.2459 ($P<0.01$) Log sum of S 0.0866 ($P<0.01$) FFM 0.0553 (NS) Between mean PD and BMI 0.1306 ($P<0.01$) WHR 0.1596 ($P<0.01$) FFM Log sum of S 0.1100 ($P<0.01$) FFM –0.0560 (NS) Between mean gingival bleeding and BMI 0.0945 ($P<0.01$) WHR 0.2081 ($P<0.01$) Log sum of S 0.0413 (NS) FFM 0.1872 ($P<0.01$)

(continued)

Table 3.8 (continued)

Author (country)	Population/sample	Body composition (exposure)	Periodontal outcome (dependent variables)	Covariates included for adjustment	Publication conclusions	Effect OR (95% CI) or other as stated
Saito et al. (2003) (Japan)	Subset of Fukuoka Health Promotion Centre health promotion program $n=643$ healthy subjects (512 females, 131 males) 19–79 years	BMI Categorized as ≤ 21.9 22–24.9 25–29.9 ≥ 30 WHR Upper body obesity defined as WHR ≥ 0.8 in females or WHR ≥ 0.9 in males BF X-ray absorptiometry (% BF) categorized For men (for women): Low ≤ 19.9 (≤ 29.9) Mid 20–29.9 (30–39.9) High 30–39.9 (40–49.9) Very high ≥ 40 (≥ 50)	10 teeth probed (1st and 2nd molars, upper right incisor, and lower left incisor) with WHO probe 6 sites per tooth giving a score/sexant Code 0 PPD < 3.5 mm Code 1 PPD 3.5–5.5 mm Code 2 PPD ≥ 6 mm Perto case classed as at least one sextant with code 1 or 2	Age, gender, smoking, socioeconomic status, diabetes, frequency of oral hygiene	Prevalence In subjects with high waist-hip ratio, higher categories of BMI (or body fat) significantly increased the adjusted risk of periodontitis, compared with subjects with low waist-hip ratios and the lowest category of BMI (or body fat)	High WHR (Upper body obesity) Risk for periodontitis (OR) 2.0 (1.4, 2.9) High WHR and BMI category Risk for periodontitis BMI 22–24.9 2.0 (1.1, 3.4) BMI 25–29.9 3.3 (1.9, 5.6) BMI ≥ 30 4.3 (1.6, 11.7)
Saito et al. (1998) (Japan)	Subset of Fukuoka Health Promotion Centre health promotion program $n=241$ healthy subjects (172 females, 69 males) 20–59 years	BMI Categorized as ≤ 20 20–24.9 25–29.9 ≥ 30 BF X-ray absorptiometry (% BF)	CPITN PPD scored one code per sextant based on deepest pocket using WHO probe CPITN Code 0–2 no periodontitis (PPD ≤ 3.5 mm) CPITN Code 3 or 4 periodontitis (PPD ≥ 4.0 mm)	Age, gender, smoking, oral hygiene frequency	Prevalence Obesity and periodontitis are related	Relative Risk for periodontitis BMI 20–24.9 1.7 (0.7, 3.8) BMI 25–29.9 3.4 (1.2, 9.6) BMI ≥ 30 8.6 (1.4, 51.4) $P=0.02$ RR for each 5% increase in body fat 1.3 (1.0, 1.6) $P=0.02$
Johansson et al. (1994) (Sweden)	Swedish subset of sample of the Multinational Monitoring of Trends and Determinants in Cardiovascular Disease study $n=1,583$ 25–64 years	BMI WHR Data collected at medical examination visit	Dental self-assessment of dentate status based on three questions: (1) Do you only have your natural teeth? (2) Do you have natural teeth and a removable partial denture? (3) Do you wear full dentures?	Age, Age, education, diet, HDL cholesterol, smoking	Prevalence Age adjusted odds ratios to be edentulous increased for, both sexes by increasing quintiles of BMI, waist hip ratio.	Risk for edentulism by increasing quintile BMI(OR): Men 1st quintile – reference 4th quintile – 1.19 (0.66, 2.14) 5th quintile – 2.07 (1.17, 3.63) Women 4th quintile – 1.85 (0.99, 3.46) 5th quintile – 3.44 (1.87, 6.32) Risk for edentulism by increasing quintile of WHR Men 1st quintile – reference 4th quintile – 1.66 (0.92, 3.00) 5th quintile – 2.38 (1.34, 4.22) Women 4th quintile – 4.40 (2.34, 8.30) 5th quintile – 5.98 (3.19, 11.2) BMI Risk to be edentulous (with analysis adjustment for all factors listed) Men 1.13 (0.94, 1.35) Women 1.34 (1.09, 1.64) Mean BMI in: Dentate men 25.8 (SE 0.1) Edentulous men 6.7 (SE 0.3) $P<0.01$ Dentate women 25.0 (SE 0.2) Edentulous women 26.8 (SE 0.3) $P<0.001$

Summers and Oberman (1962) (USA)	Subset of Tecumseh, Michigan health study 408 participants had dental exam 84 edentulous excluded $n = 3,24 >20$ years Males = 154 Females = 170	Weight Relative weight	Periodontal Disease Index (PDI) Based on examination of Ramjord teeth (UR molar, UL central, UL bicuspid, LL molar, LR central, LL bicuspid Categorical score given assessing periodontal status based on gingival inflammation, crevice depth, plaque and calculus	Genders analysed separately	Severity The relationship between periodontal disease and certain physiologic and social variables might be worthy of longitudinal study	Correlation (simple correlation coefficient): between weight and PDI score in men -0.78 (NS) between weight and PDI score in females 0.13 (NS) between relative weight and PDI score in men 0.82 (NS) between weight and PDI score in females -0.32 (NS)
----------------------------------	--	------------------------	---	-----------------------------	--	--

BF body fat, *BMI* body mass index, *BOP* bleeding on probing, *CAL* clinical attachment level, *CI* confidence interval, *CPITN* Community Periodontal Index of Treatment Needs, *GB* gingival bleeding index, *N/A* no data, *NHANES III* Third National Health and Nutrition Examination Survey, *OR* odds ratio, *PD* pocket depth, *PI* plaque index, *PPD* pocket probing depth, *RR* relative risk, *SD* standard deviation, *SE* standard error, *VFA* visceral fat area, *WC* waist circumference, *WHO* World Health Organization, *WHR* waist/hip ratio

Fig. 3.7 Case-control study design (Morabia 1997. Reprinted with permission from Elsevier)



Fig. 3.8 Cohort study design (Morabia 1997. Reprinted with permission from Elsevier)

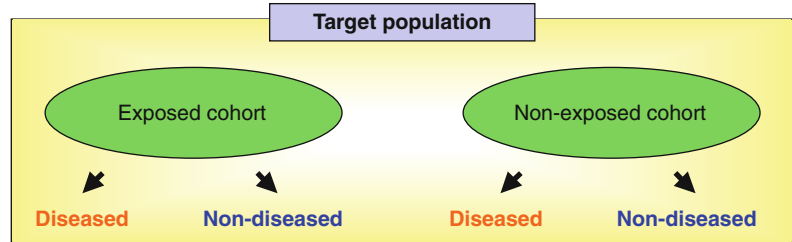
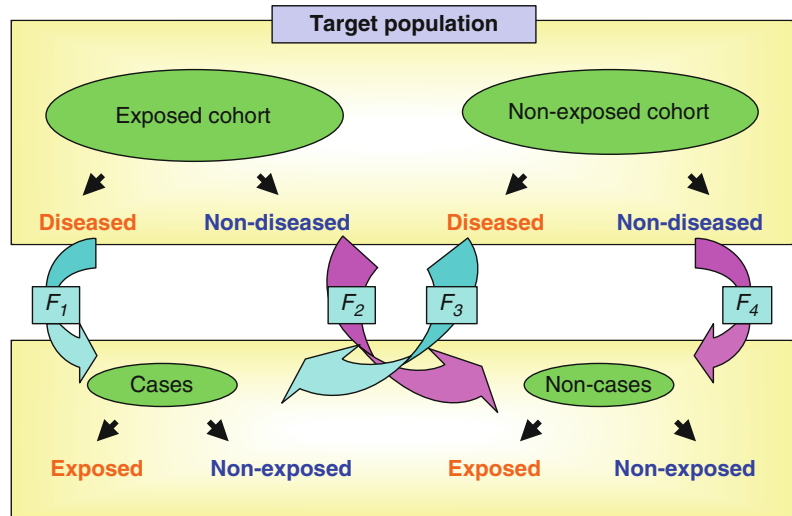


Fig. 3.9 Relation of the case-control study design (lower box corresponding to Fig. 3.7) to the cohort study design (upper box, corresponding to Fig. 3.8). F_4 sampling fraction. Dotted lines of the arrows corresponding to F_2 and F_4 indicate that in population-based case-control studies F_1 and F_3 will, in general, be much greater than F_2 and F_4 (Morabia 1997. Reprinted with permission from Elsevier)



The Achilles heel of case-control studies is choosing an **appropriate control group** (Grimes and Schulz 2002a). Choosing controls might seem deceptively simple, but it can be treacherous. Controls should reflect the background frequency of the exposure in the population. Hence, they should be similar in all important respects to cases, except that they do not have the disorder in question. Their selection must be independent of exposure (Grimes and Schulz 2005).

Controls can be taken from known or unknown study populations. A known group consists of a defined population observed over a

period. When the study group of potential controls is known, a good approach is to take all or, if not feasible, a random sample of them. If no population roster exists, then techniques such as random digit dialling can be used (Fig. 3.10) (Grimes and Schulz 2005). When the group of potential controls is unknown, choosing controls gets tough. Generally, it is recommended to use individuals chosen from the same time and place as cases. In this situation, hospital controls, neighbourhood controls, and friend, associate, or relative controls can be used (Fig. 3.11) (Grimes and Schulz 2005). If the

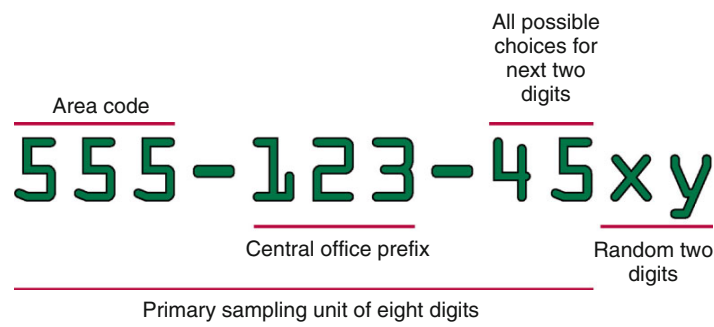


Fig. 3.10 Random-digit dialling for controls. Primary sampling unit included eight-digit numbers: all known area codes and three-digit central-office prefixes in the county, plus all combinations of next two digits. For all

these eight-digit numbers randomly chosen, a computer generated the two final digits, creating a ten-digit number to be called (Reprinted from Grimes and Schulz (2005), with permission from Elsevier)

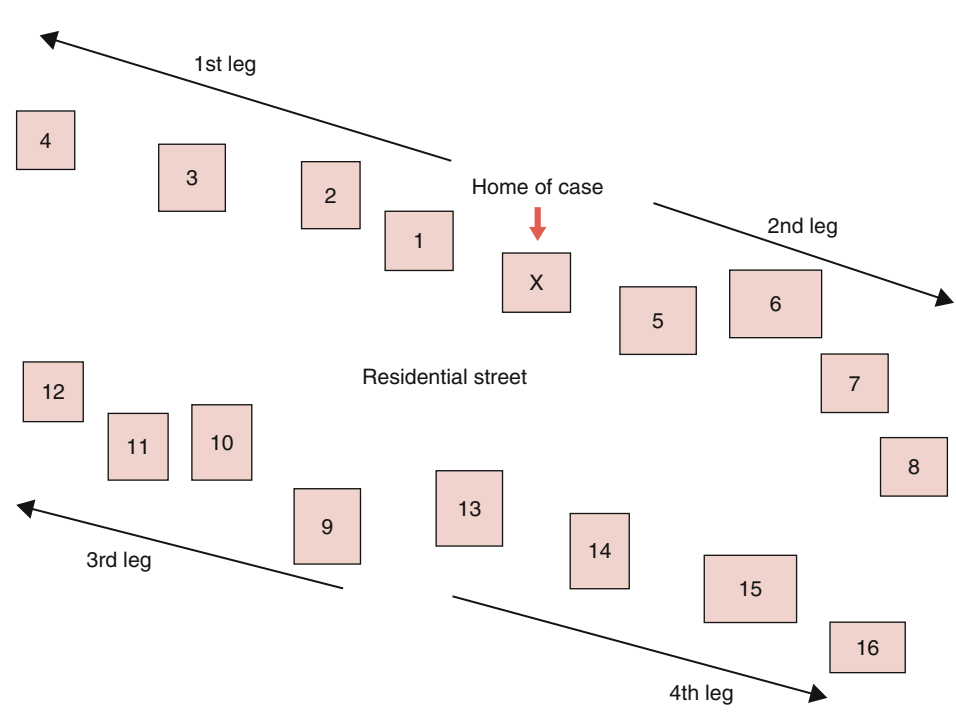


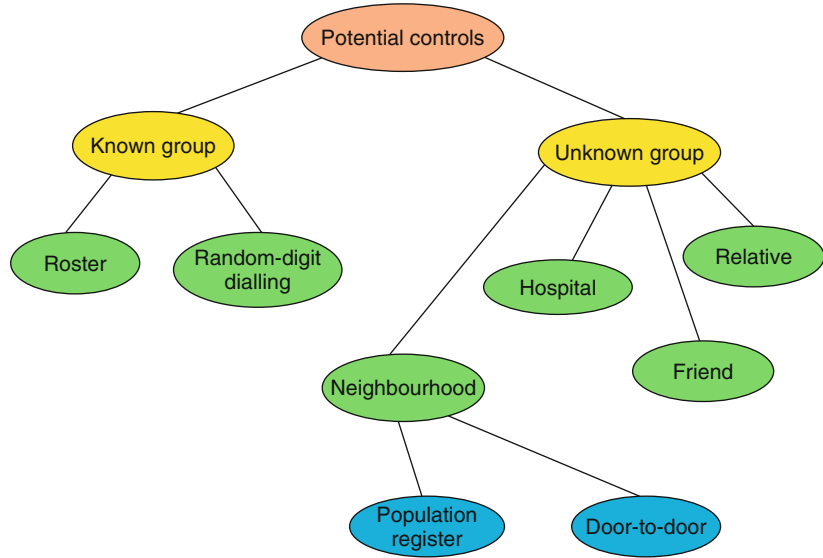
Fig. 3.11 Neighbourhood controls. After interviewing the case in her home, the investigator canvasses up to 16 homes in a predetermined H-shaped pattern until three controls are identified and interviewed. *Every rectangle*

represents a home along the same street (Reprinted from Grimes and Schulz (2005), with permission from Elsevier; Olson et al. 1992. Reprinted with permission from Oxford University Press)

appropriateness of a control group is uncertain, sometimes a second control group is added. If the number of cases is small, having up to four times as many controls improves study power. However, this strategy does not improve validity (**Fig. 3.12**) (Grimes and Schulz 2005).

Another difficulty in case-control studies involves the measurement of exposure information. Participants, both cases and controls, might inaccurately remember past exposures, especially those that happened a long time ago. Furthermore, cases often remember exposures to putative risk

Fig. 3.12 Choosing controls with known and unknown group of study participants (Reprinted from Grimes and Schulz (2005), with permission from Elsevier)



factors differently than controls. This differential recall (**recall bias**) causes information bias (Schulz and Grimes 2002).

Because the case–control study lacks denominators, investigators cannot calculate incidence rates, relative risks, or attributable risks. Instead, **odds ratios** are the measure of association used; when the outcome is uncommon, the odds ratio provides a good proxy for the true relative risk (Grimes and Schulz 2002a).

Five main notions guide investigators who do, or readers who assess, case–control studies:

1. First, investigators must explicitly define the criteria for diagnosis of a case and any eligibility criteria used for selection.
2. Second, controls should come from the same population as the cases, and their selection should be independent of the exposures of interest.
3. Third, investigators should blind the data gatherers to the case or control status of participants or, if impossible, at least blind them to the main hypothesis of the study.
4. Fourth, data gatherers need to be thoroughly trained to elicit exposure in a similar manner from cases and controls; they should use memory aids to facilitate and balance recall between cases and controls.
5. Finally, investigators should address confounding in case–control studies, either in the design stage or with analytical techniques. Devotion of meticulous attention to these points enhances the validity of the results and bolsters the reader’s confidence in the findings (Schulz and Grimes 2002).

Examples

Table 3.9 gives details of case–control studies analysed in a meta-analysis of association of periodontitis with increased risk of coronary heart disease (Bahekar et al. 2007). The case–control studies (1,423 patients) showed a great risk of developing CHD ($OR = 2.22$, 95% CI 1.59–3.117, $P < 0.001$).

3.3.6 Cohort Studies

The term cohort has military, not medical, roots. A cohort was a 300–600-man unit in the Roman army; ten cohorts formed a legion (Fig. 3.13). The etymology of the term provides a useful mnemonic: a cohort study consists of bands or groups of persons marching forward in time from an exposure to one or more outcomes (Grimes and Schulz 2002d).

Table 3.9 Details of case-control studies included in the meta-analysis of association of periodontitis with increased risk of coronary heart disease (Bahekar et al. 2007) (Reprinted with permission from Elsevier)

Reference	Study population	Assessment of oral health	Cardiovascular assessment	Variables controlled	Quality rating
Malthaner et al. (2002)	Nonsmoking and nondiabetic patients older than 40 years with no history of MI in past 6 months and who had undergone cardiac catheterization in past 12 months	Periodontal examination evaluating clinical attachment level and also radiographic analysis assessing radiographic bone level	Angiographically confirmed CAD. Cases: 50% stenosis in at least one major epicardial artery. Controls: <50% stenosis in all identified arteries	Adjusted for age and previous smoking status	Fair
Geerts et al. (2004)	108 CAD patients and 62 healthy controls	Probing depth, periodontal pocket bleeding index, plaque index, and tooth mobility	Cases: 108 patients treated for CAD. 62 healthy controls with similar mean age group	Age, sex, smoking, alcohol intake, diet, physical activity, HTN, diabetes, hyperlipidemia	Good
Buhlin et al. (2005)	143 women aged 43–79 years with CHD and 50 women aged 45–77 years without CHD	Dental examination measuring probing depth score, hygiene index, restoration, and dental caries. Digital panoramic radiographs measuring bone height	Cases: CHD patients who were treated for MI, PTCA, and CABG in two large hospitals. Controls: women with no history of CHD	Age, smoking, BMI, diabetes, education, and place of birth	Good
Briggs et al. (2006)	Cases were men aged 40 years and older with angiographically proven CHD, and controls were age-matched men.	Measurement of plaque, bleeding on probing, probing depth measurement	Cases were men with proven CAD confirmed by N50% narrowing of at least one coronary artery at coronary angiography. Controls were the random sample of age-matched men from the same locality	Smoking, academic achievement, alcohol consumption, unemployment, ability to maintain the body weight, regular exercise, ability to relax daily, having a hobby or pastime, plaque, and C-reactive protein	Good
Spahr et al. (2006)	263 angiographically confirmed CHD subjects and 526 age and sex-matched controls without history of CHD	A modified CPITN measured at six sites of each tooth	Cases had at least one stenosis that was 50% or more of the luminal diameter of a major coronary artery	Age, sex, BMI, smoking, alcohol consumption, DM, HTN, hyperlipoproteinemia, level of education, physical activity, and statin intake	Good

PTCA percutaneous transluminal coronary angioplasty, CPITN Community Periodontal Index of Treatment Needs, CAD coronary artery disease, MI myocardial infarction, CABG coronary artery bypass graft, DM diabetes mellitus, HTN hypertension

Fig. 3.13 An early cohort in search of favourable outcomes (Reprinted from Grimes and Schulz (2002d), with permission from Elsevier)

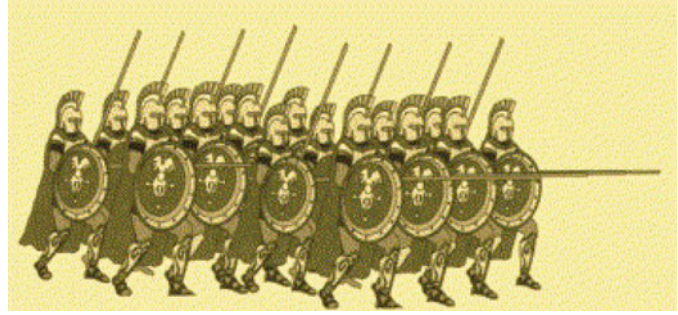
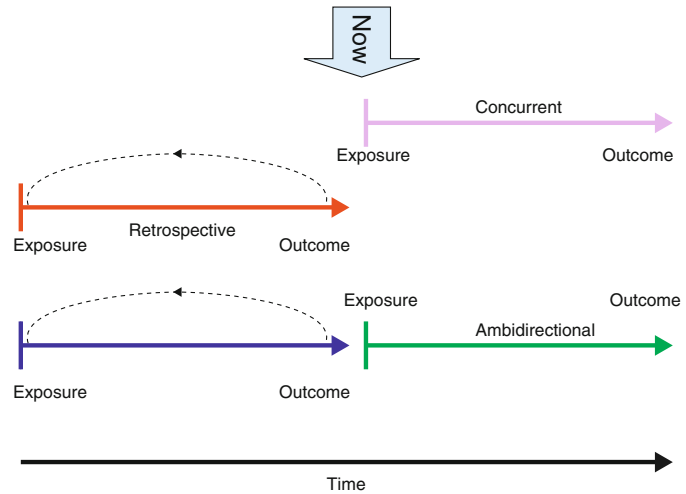


Fig. 3.14 Schematic diagram of concurrent, retrospective, and ambidirectional cohort studies (Grimes and Schulz 2002d), (Reprinted from The Lancet 2002;359:341–345 with permission from Elsevier)



This analogy might be helpful since cohort studies have a bevy of confusing synonyms: **incidence, longitudinal, forward-looking, follow-up, concurrent, and prospective study**. Although the terminology can seem daunting, the cohort study is easy for clinicians to understand since it flows in a logical direction (unlike the case–control study) (Grimes and Schulz 2002d). Cohort studies proceed in a logical sequence: from exposure to outcome. Hence, this type of research is easier to understand than case–control studies. Investigators identify a group with an exposure of interest and another group or groups without the exposure. The investigators then follow the exposed and unexposed groups forward in time to determine outcomes. If the exposed group develops a higher incidence of the outcome

than the unexposed, then the exposure is associated with an increased risk of the outcome (Grimes and Schulz 2002a).

While all cohort studies run forward in time from exposure to outcome, not all are done in real-time (**Fig. 3.14**). Stated alternatively, cohort studies can be **concurrent, non-concurrent, or ambidirectional**. A **concurrent** cohort study enrolls persons exposed and unexposed and follows them forward contemporaneously to determine who gets the outcome of interest. A **non-concurrent** cohort study goes back in time to comprise the groups of exposed or unexposed (e.g. through hospital charts or employment records) and tracks them forward in time to determine outcomes. Although the forward direction is the same, the data collection process is not contemporaneous.

An **ambidirectional** cohort study combines both approaches above; data gathering is done historically and contemporaneously. This can expedite getting results (Grimes and Schulz 2002c).

What to look for in cohort studies? (Grimes and Schulz 2002d):

1. **Who is at risk?** All participants (both exposed and unexposed) in a cohort study must be at risk of developing the outcome.
2. **Who is exposed?** Cohort studies need a clear, unambiguous definition of the exposure at the outset. This definition sometimes involves quantifying the exposure by degree, rather than just yes or no. For example, the minimum exposure might have to be 14 cigarettes/day or less. Definition of exposure levels in this way can result in more than two groups—for example, non-smokers, light smokers, and heavy smokers.
3. **Who is an appropriate control?** The key notion is that controls (the unexposed) should be similar to the exposed in all important respects, except for the lack of exposure. If so, the unexposed group will reveal the background rate of the outcome in the community. The unexposed group can come from either internal (persons from the same time and place, such as a hospital ward) or external sources. Internal comparisons are most desirable. In a particular population, individuals segregate by themselves (or through medical interventions) into exposure status—for example, cigarette smoking, occupation, and contraception
4. **Have outcomes been assessed equally?** Outcomes must be defined in advance; they should be clear, specific, and measurable. Identification of outcomes should be comparable in every way for the exposed and unexposed to avoid information bias. Failure to define objective outcomes leads to uninterpretable results.

How to run a cohort study? If the data are readily available, then a retrospective design is the quickest method. If high-quality, reliable data are not available, a prospective study will be required. The first step is the definition of the sample group. Each subject must have the potential to develop the outcome of interest. Furthermore, the sample population must be representative of the general

population if the study is primarily looking at the incidence and natural history of the condition (descriptive). If however the aim is to analyse the relation between predictor variables and outcomes (analytical), then the sample should contain as many patients likely to develop the outcome as possible; otherwise, much time and expense will be spent collecting information of little value. Each variable studied must be accurately measured. To minimize the potential for missing a confounding variable, all probable relevant variables should be measured. If this is not done, the study conclusions can be readily criticized. All patients entered into the study should also be followed up for the duration of the study. Losses can significantly affect the validity of the results. To minimize this, as much information about the patient (name, address, telephone, GP, etc.) needs to be recorded as soon as the patient is entered into the study. Regular contact should be made; it is hardly surprising if the subjects have moved or lost interest and become lost to follow up if they are only contacted at 10-year intervals! (Mann 2003).

In population-based cohort studies, a sample, or even the entirety, of a defined population is selected for longitudinal assessment of exposure–outcome relations. Because of their typically high cost and logistic complexities, population-based cohort studies generally evaluate **multiple hypotheses**, some defined a priori as well as some suggested in the course of the study either based on interim analyses or on advances in the field, particularly when data are repeatedly collected on cohort members. Perhaps the most often cited justification for conducting a population-based study is its external validity—that is, the applicability of its results to a defined population (Szklo 1998).

Advantages of cohort studies. Cohort studies allow direct determination of relative risk. They provide strong evidence of a disease relationship and permit the calculation of lag times between exposure or risk factor and disease. The cohort study enables calculation of true incidence rates, relative risks, and attributable risks (Grimes and Schulz 2002a). Other outcome measures in cohort studies include life table rates, survival curves, and hazard ratios (Table 3.10). By contrast, case–control studies

Table 3.10 Reporting time-to-event in cohort studies (Grimes and Schulz 2002d) (Reprinted from The Lancet 2002;359:341–345 with permission from Elsevier)

Outcome measures	Description	Examples
Survival analysis	Survival analysis is useful when lengths of follow-up vary substantially or when participants enter a study at different times. The Kaplan-Meier method provides a more sophisticated expression of the risk of the outcome over time than does a simple dichotomous outcome. It can determine the probability (<i>P</i>) of the outcome at any point in time; this result is graphed as a step function (which jumps at every event). A complementary, mirror-image graph portrays the likelihood of avoiding the outcome ($1 - P$) as a function of time (Kaplan-Meier survival curve). The log-rank test compares survival curves of different groups (Grimes and Schulz 2002d).	Figure 3.15 depicts the long-term outcomes of implants placed in patients treated for periodontitis periodontally compromised patients and in periodontally healthy patients in relation to adherence to supportive periodontal therapy. To evaluate the implant survival rates in the three groups of patients, the Kaplan-Meier analysis with log-rank pooled per strata was adopted both for all the implants and for the solid screws only. The survival rate was 96.6%, 92.8% and 90% for all implants and 98%, 94.2% and 90% for the solid-screw implants, respectively, for periodontally healthy patients, moderate periodontitis periodontally compromised patients and severe periodontitis periodontally compromised patients (Roccuzzo et al. 2010).
Proportional hazard model	Another approach to different lengths of follow-up is the Cox proportional hazard model. It is a multivariate technique that has time-to-event (such as illness) as the dependent variable. By contrast, multiple logistic regression has 'yes-no' as the dependent variable. Coefficients from this model can be used to calculate the risk ratio (hazard ratio) of the outcome, after controlling for other covariates in the equation. The hazard ratio (with 95% CIs) is interpreted in the same way as a relative risk for dichotomous outcomes (Grimes and Schulz 2002d).	Levin et al. (2011) evaluated the long-term survival rates of dental implants according to the patient's periodontal status and estimated if the effect of periodontal status regarding implant failure is constant throughout the long-term follow-up in a historical prospective cohort study design of all consecutive patients operated from 1996 to 2006 at a periodontal clinic. The cohort consisted of 736 patients, with a total of 2,336 dental implants. The main assumption of the Cox regression is the proportional hazard (PH) assumption, which states that the HR is constant throughout the follow-up time. The results of an initial naive Cox regression model (main effects without interaction terms) indicate a violation of the PH assumption for smoking and periodontal group. This result means that smoking and periodontal status effects (measured by HR) are not constant throughout the follow-up period and therefore the PH assumption is not true. Diagnostic plots for severe chronic periodontitis and smoking effects against survival time are displayed in Fig. 3.16 . The positive slopes (non-horizontal line) seen in both figures support the finding of PH violation. According to these figures, the HR for both severe chronic and smoking are not constant with greater HR at a longer follow-up time. In order to overcome this violation, the extended Cox PH model was constructed by including an interaction term between survival time (shorter or longer than 50 months) and periodontal status as well as interaction between survival time and smoking status. The extended Cox model revealed that severe chronic status turned out to be a significant risk factor for implant failure after 50 months of follow-up [hazard ratio (HR) = 8.06; $P < 0.01$]. The extended Cox model for smoking indicates a near-significant effect after 50 months (HR = 2.76; $P = 0.061$). Notice the constant effect obtained from the extended model, seen as a horizontal line in Fig. 3.17 . It was possible to conclude that periodontal status and smoking are significant risk factors for late implant failures. The HR for periodontal and smoking status were not constant throughout the follow-up period. After 50 months, the hazard for implant failure is eight times greater for the severe periodontitis patients (Levin et al. 2011).

Fig. 3.15 The Kaplan–Meier estimate of the survival rate of solid-screw implants as a function of time since insertion in PHP (periodontally healthy patients), moderate PCP and severe PCP (periodontally compromised patients) (Roccuzzo et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

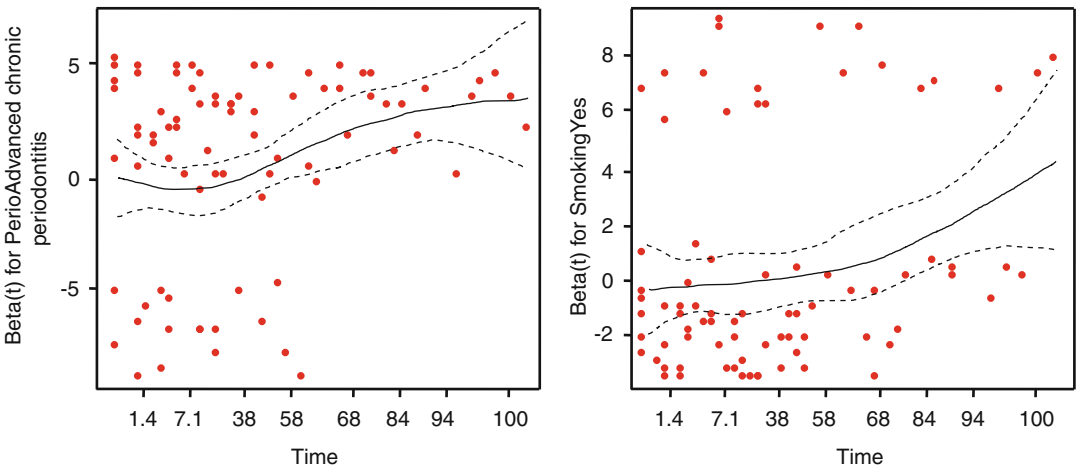
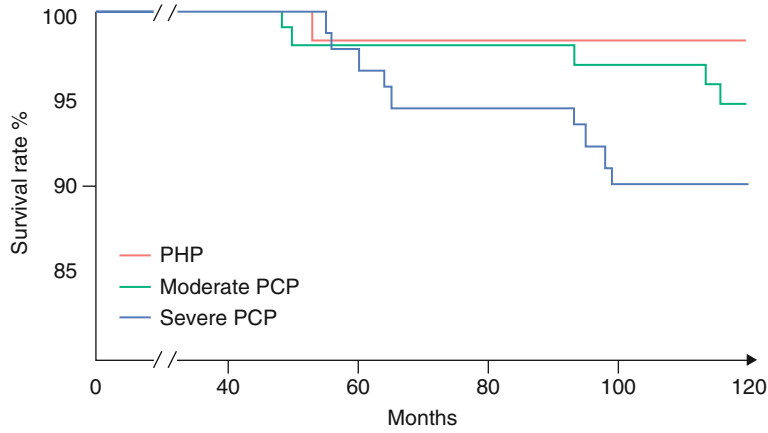
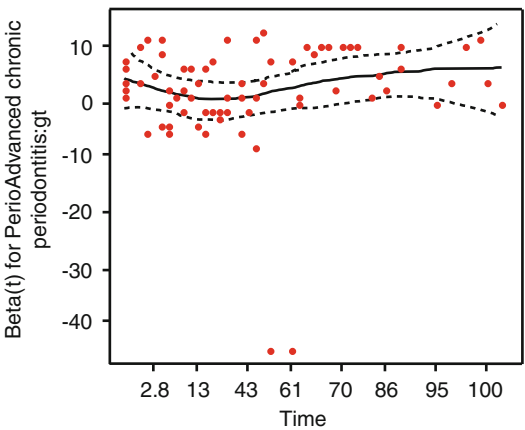


Fig. 3.16 Effect of advanced chronic periodontitis (*left panel*) and smoking (*right panel*) against survival time by fitting a naïve Cox regression model. The effect (Beta) is assumed to be constant in a Cox regression model (PH assumption). This should be reflected by a *horizontal line*.

A *non-horizontal line* indicates a violation of the PH assumption and consequently a questionable validity for the naïve Cox regression model (Levin et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

Fig. 3.17 Effect of advanced chronic periodontitis against survival time by fitting an extended Cox regression model. The extended Cox model include an interaction term between survival time (shorter or longer than 50 months) and periodontal status (Levin et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)



cannot provide incidence rates; at best, odds ratios approximate relative risks only when the outcome is uncommon (Grimes and Schulz 2002d).

The study of a cohort which is representative of a defined population offers three additional advantages: (1) It allows the estimation of distributions and prevalence rates of relevant variables in the reference population (information on risk factors is used for the calculation of population attributable risks). (2) Risk factor distributions measured at baseline in a study involving periodic examinations of the cohort can be compared with distributions in future cross-sectional samples so as to assess risk factor trends over time. (3) A representative sample is the ideal setting in which to carry out unbiased evaluations of relations, not only of confounders to exposures and outcomes but also among any other variables of interest, even those which were not specified in the original study hypotheses (Szklo 1998).

The weakness of cohort studies is found in the long duration of follow-up necessary for adequate disease endpoints to be accumulated. They are costly to perform, and loss to follow-up is a serious problem. Like cross-sectional studies, cohort studies are generally a poor choice for rare diseases (Luepker 2005).

Variations on the cohort theme:

- **Before–after studies (time series)** have important limitations. Here, an investigator takes a measurement, exposes participants to an intervention (often a drug), repeats the measurements, then compares them. First, regression to the mean is often ignored. If admission to the cohort includes extreme measurements, such as high laboratory values, then lower mean values will arise at follow-up, irrespective of treatment. Second, secular trends, such as seasonal changes in the frequency of pneumonia, can affect results. Third, washout periods are often needed to avoid a carry-over effect of drugs given during the initial observation period (Grimes and Schulz 2002d).
- Cohort studies sometimes spawn other studies. One of the most frequent is the **nested case–control study**. Why would an investigator carve out a case–control study in the midst of a cohort study? Some exposure or predictor

variables are simply too expensive to determine on everyone in a study. A clever way to skirt this financial obstacle is to do a cohort study that will yield a sufficient number of cases. All those in the cohort study who develop the outcome of interest now become the cases for the nested study. The investigator then chooses a random sample of all participants who did not develop the outcome (controls). Controls are generally matched to cases by important characteristics, such as age and sex (Grimes and Schulz 2002d).

For example, a recent nested case–control study examined the relationship between maternal periodontal disease and plasma angiogenic factor expression of soluble fms-like tyrosine kinase (sFlt)-1 (Horton et al. 2009). A total of 1,020 women were enrolled in the prospective, observational Oral Conditions and Pregnancy (OCAP) study and considered for this analysis. The OCAP study was a prospective cohort study of maternal periodontal disease and obstetric outcomes performed at the University of North Carolina School of Dentistry Center for Oral and Systemic Disease, in collaboration with Duke University Medical Center, Durham, North Carolina, from December 1997 through July 2001 (Leiff et al. 2004). Of these women, 220 had oral health examination, and plasma soluble fms-like tyrosine kinase (sFlt)-1, sFlt-1, also known as soluble vascular receptor endothelial growth factor receptor-1, results were available and were included in this analysis. Fifty five women without active periodontal disease were compared to 17 women with moderate/severe periodontal infection, 43 women without evidence of periodontal disease progression were compared to 28 women with periodontal disease progression, and 64 women without evidence of foetal exposure to oral pathogens were compared to 13 women with foetal exposure to oral pathogens. The median sFlt-1 concentration at the time of enrolment for all women was 2,374 pg/ml (interquartile range [IQR] 1,504–3,194 pg/ml). Women with evidence of foetal exposure to oral pathogens had significantly higher sFlt-1 concentrations compared to IgM-negative fetuses (3,383 pg/ml [IQR 2,610–4,244 pg/ml] versus

2,123 pg/ml [IQR 1,456–3,011 pg/ml]; $P=0.03$). Thus, instead of measuring sFlt-1 concentrations of the entire cohort of 1,020, the investigators incurred this laboratory expense for 20% of the cohort (Horton et al. 2009).

Examples

Table 3.11 includes the description of cohort studies included in a recent meta-analysis of periodontal disease and risk of coronary heart disease and stroke (Janket et al. 2003).

3.3.7 Other Design Types

A **cost–benefit analysis** is an additional type of research study that does not fit the usual classification systems. These studies are a form of economic assessment in which the costs of medical care are compared with the economic benefits of that care. The benefits generally include a calculation for increased earnings due to improved health, as well as potential reductions in future health-care costs. Generally, these calculations are done from a societal perspective. A number of assumptions are always built into these analysis models. For example, what is the value of an additional year of life? The authors should clearly identify the nature, range, and reasons for any assumptions (Thompson and Panacek 2007).

The disadvantage of cost–benefit analyses is that a validation in monetary units of the clinical outcome has to be presented, that is, the monetary value of a surviving tooth and or reconstruction or the money saved by avoiding therapy (Braegger et al. 2005).

A **cost-effectiveness analysis** is similar but compares alternative programmes, therapies, or other interventions in terms of their overall costs per degree of clinical effect (Thompson and Panacek 2007).

This type of economic evaluation balances costs in monetary units against the outcome presented in non-monetary units, that is, change in clinical parameters, survival, prevented surgical interventions, and reduction of tooth mortality (Fig. 3.18). It is important to clearly separate the

term efficacy, defined as the probability of a service being beneficial to patients provided under ideal conditions, from the term effectiveness, which concerns the care provided to the general population under conditions usually found in practice. A disadvantage of cost-effectiveness analyses is that only interventions with the same clinical endpoint can be compared. In reality, however, the clinical outcome may vary considerably (survival of a tooth with/without recession, mobility, sensitivity, aesthetic impairment, etc.) (Braegger et al. 2005).

Although the scientific approaches used in economic evaluation analyses are becoming more sophisticated, this methodology is not yet standardized and has tremendous variability in the methods used within different studies. This type of research generally requires a team effort, including panels of clinicians, biostatisticians, and, sometimes, economists. The work itself is often quite tedious and is not a form of research for the novice (Thompson and Panacek 2007).

Gjerme and Grytten (2009) assessed the cost-effectiveness of various modalities for chronic periodontitis in adults. They searched the literature for studies where time consumption and monetary costs (where available) were used as expressions of the resources required to perform the following types of services: providing information on the disease and oral hygiene instruction for patients, removal of sub-gingival calculus, access surgery, and maintenance and prevention of recurrence of the disease. The following outcome variables were chosen: retention of teeth or a surrogate variable such as change in bone level measured on X-rays taken at the start and at the end of a 1-year (or longer) period. The literature search did not reveal any studies where these outcome variables could be linked to information on resources used in treatment. This review revealed that in most cases, it was difficult to assess whether one type of treatment (or intervention) is more effective than another. It was concluded that better data are needed, in particular on outcomes, so that valid comparisons can be made between different types of treatment, for different types of studies (Gjerme and Grytten 2009).

Table 3.11 Description of cohort studies included in a meta-analyses of periodontal disease and risk of coronary heart disease and stroke (Janket et al. 2003) (Reprinted with permission from Elsevier)

Authors	Study type/ quality score	Length of follow-up (y)	Sample size no. of cases	Relative risk (CI)	Age category (y)	Outcome measurement	Predictor measurement	Confounders adjusted
Beck et al. (1996)	Prospective cohort ^a	18	1,147/207	1.49 (1.07–2.15)	21–80	Fatal MI, nonfatal MI, stroke (defined by ICD-9 code 410–414)	With probing pocket depth, alveolar bone loss	Diabetes, BMI, SBP, cholesterol
DeStefano et al. (1993)	Prospective cohort ^b	14	9,760/1,425	1.25 (1.06–1.48)	25–75	Mortality d/t CHD, admission d/t CHD (defined by ICD-9 code 410–414)	Gingivitis, mild periodontitis, moderate periodontitis, severe periodontitis	Race, education, SBP, cholesterol, diabetes, alcohol, smoking, BMI, exercise, poverty
Genco et al. (1997)	Prospective cohort ^a	10	1,372/68	2.68 (1.30–5.50)	<60	CVD with EKG	Tooth loss, alveolar bone loss	Diabetes, cholesterol, hypertension
Howell et al. (2001)	Prospective cohort ^c	12.5	22,037/2,042	1.01 (0.86–1.15)	40–84	Death due to CVD, nonfatal MI, nonfatal stroke	History of periodontitis	Diabetes, hypertension, BMI, smoking, exercise, alcohol use
Hujoel et al. (2000)	Prospective cohort ^b	8–10	8,032/1,265	1.14 (0.96–1.36)	25–74	First hospitalization, revascularization, death from CHD	No periodontitis, gingivitis, periodontitis	Race, education, poverty, marital status, SBP, DBP, total cholesterol, diabetes, exercise, BMI, alcohol use, smoking, nervous breakdown
Joshiyura et al. (1996)	Prospective cohort ^a	6	44,119/757	1.04 (0.86–1.25)	<60 (male)	Fatal MI, nonfatal MI	questionnaire/history of periodontitis	Smoking, BMI, diet habit, exercise, family MI history
Mattila et al. (1995)	Prospective cohort ^b	7.2	214/52	1.21 (1.08–1.36)	<65 (female)	Death from CVD, admission for CVD	Total dental index, including periodontitis, periapical lesion, dental caries, pericoronitis	Diabetes, BMI, hypertension, smoking, cholesterol, triglyceride, SES
Morrison et al. (1999)	Retrospective cohort ^c	21	10,368/416	2.15 (1.25–3.72)	35–84	Fatal CHD	No periodontitis, mild gingivitis, severe gingivitis, periodontal-pocket	Cholesterol, smoking, diabetes, hypertension, residence
Wu et al. (1999)	Prospective cohort ^b	10+	9,962/803	1.17 (1.04–1.31)	25–74	Fatal and nonfatal strokes (including hemorrhagic, nonhemorrhagic, and transient ischemic events)	No periodontal disease, gingivitis, gingivitis with pocket, advanced periodontitis	Race, education, poverty index, smoking, diabetes hypertension, alcohol use cholesterol, BMI

Janket et al. (2003). Reprinted with permission from Elsevier

All studies adjusted for age and sex

MI myocardial infarction, ICD International Classification of Diseases, EKG electrocardiogram, BMI body mass index, SBP systolic blood pressure, DBP diastolic blood pressure, SES socioeconomic status

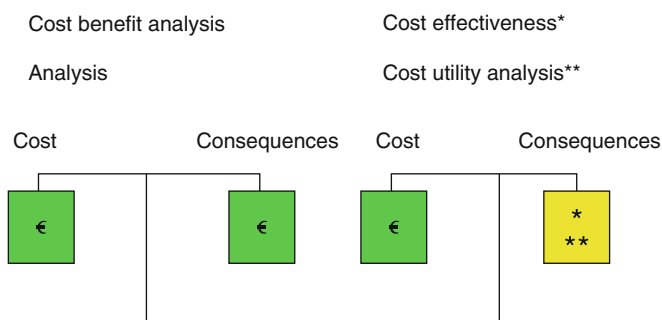
Quality score:

^a45–59% of maximum score achieved

^b80–90% of maximum score achieved

^c60–79% of maximum score achieved

Fig. 3.18 Economic analyses, with input and output. *Change in clinical parameters or survival, years of life, number of retained teeth or **quality adjusted life years of teeth/dentition/reconstruction (Braegger 2005. Reprinted with permission from John Wiley & Sons. Inc.)



The original, benchmark analysis of the cost-effectiveness of treatment options for periodontal disease applied a measure of utility analogous to the quality-adjusted life year—the quality-adjusted tooth year; the value of teeth with sensitivity or poor aesthetics relative to teeth with no sensitivity or good aesthetics was estimated from standard gamble exercises with a sample of patients attending for periodontal treatment. The cost-effectiveness of the following different treatment options was explored: no periodontal treatment (6-monthly prophylaxis), non-surgical management, surgical treatment with and without osseous recontouring, and non-surgical management with systemically administered antimicrobial therapy. Estimates of tooth life expectancy and the utility values of healthy and diseased teeth were combined in a model to determine an estimated number of quality-adjusted tooth years for each treatment option. The effectiveness of the different treatments was determined from a literature review of 24 clinical trials, and rates of tooth loss following the treatment options were estimated by a panel of periodontists. An extensive series of evaluations was completed for eight different disease severities, and single-rooted and multiple-rooted teeth were considered for each state (16 different scenarios in total) (Heasman et al. 2011).

The cost-effectiveness analyses, which were considered to be groundbreaking at the time, presented a series of incremental cost-effectiveness ratios for the treatment options that were superior to 'no treatment' (6-monthly prophylaxis) for teeth affected by varying degrees of periodontal disease severity. The lowest incremental cost-

effectiveness ratios were produced when systemic antimicrobial therapy plus non-surgical treatment was compared with non-surgical treatment alone (\$1.36–\$20.00 per quality-adjusted tooth year in 1987 terms). An economic analysis indicates that systemic antimicrobials are more cost-effective than locally delivered antimicrobial drugs for the treatment of moderate or advanced periodontitis. The analysis, however, has not considered the increased risk of bacterial resistance that might influence patient management, particularly so by compromising successive episodes of periodontal treatment with antimicrobials. It was concluded that the cost-effectiveness of the different preparations beyond 12 months is somewhat speculative and there is no confidence that short-term data are in any way predictive of long-term outcomes. The paucity of long-term clinical outcomes demands longer-term studies to demonstrate whether the effects of antimicrobials on clinical attachment loss in the short term relate ultimately with advancement of tooth loss or effects on any other patient-centred outcomes (Heasman et al. 2011).

3.3.8 The STROBE Statement

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement is a checklist of items that should be addressed in articles reporting on the three main study designs of analytical epidemiology: cohort, case-control, and cross-sectional studies. The intention is solely to provide guidance on how to report observational research well: These

recommendations are not prescriptions for designing or conducting studies. Also, while clarity of reporting is a prerequisite to evaluation, the checklist is not an instrument to evaluate the quality of observational research (von Elm et al. 2007).

The STROBE Statement is a checklist of 22 items that we consider essential for good reporting of observational studies. These items relate to the article's title and abstract (item 1), the introduction (items 2 and 3), methods (items 4–12), results (items 13–17) and discussion sections (items 18–21), and other information (item 22 on funding). Eighteen items are common to all three designs, while four (items 6, 12, 14, and 15) are design specific, with different versions for all or part of the item (**Table 3.12**). For some items (indicated by asterisks), information should be given separately for cases and controls in case-control studies or exposed and unexposed groups in cohort and cross-sectional studies. Although presented here as a single checklist, separate checklists are available for each of the three study designs on the STROBE website (<http://www.strobe-statement.org/index.php?id=available-checklists>) (von Elm et al. 2007).

The STROBE Statement was not developed as a tool for assessing the quality of published observational research. Such instruments have been developed by other groups and were the subject of a recent systematic review (Sanderson et al. 2007). In the Explanation and Elaboration paper, several examples of good reporting from studies whose results were not confirmed in further research are used—the important feature was the good reporting, not whether the research was of good quality (Vandenbroucke et al. 2007). However, if STROBE is adopted by authors and journals, issues such as confounding, bias, and generalizability could become more transparent, which might help temper the overenthusiastic reporting of new findings in the scientific community and popular media, and improve the methodology of studies in the long term. Better reporting may also help to have more informed decisions about when new studies are needed and what they should address (von Elm et al. 2007).

3.4 Experimental Studies

3.4.1 Non-randomized Trial

Some experimental trials do not randomly allocate participants to exposures—for example, treatments or prevention strategies. Instead of using truly random techniques, investigators often use methods that fall short of the mark—for example, alternate assignment (Grimes and Schulz 2002a).

After the investigators have assigned participants to treatment groups, the way a non-randomized trial is done and analysed resembles that of a cohort study. The exposed and unexposed are followed forward in time to ascertain the frequency of outcomes. Advantages of a non-randomized trial include use of a concurrent control group and uniform ascertainment of outcomes for both groups. However, selection bias can occur (Grimes and Schulz 2002a).

Kunz and Oxman (1998) realized comparisons of randomized clinical trials and non-randomized clinical trials, trials with adequately concealed random allocation versus inadequately concealed random allocation. **Table 3.13** summarizes the eight studies comparing randomized clinical trials and nonrandomized clinical trials of the same intervention. In five of the eight studies, estimates of effect were larger in non-randomized trials. Outcomes in the randomized treatment groups and nonrandomized treatment groups were frequently similar, but worse outcomes among historical controls spuriously increased the estimated treatment effects. One study found comparable results for both allocation procedures, and two studies reported smaller treatment effects in nonrandomized studies. In one study, the smaller estimate of effect was due to a poorer prognosis for patients in the non-randomized treatment groups. The deviation of the estimates of effect for nonrandomized trials compared with randomized trials ranged from an underestimation of effect of 76% to an overestimation of effect of 160%. Several explanations for discrepancies between estimates of effect derived from randomized trials and non-randomized trials are possible. For example, it

Table 3.12 The STROBE statement – checklist of items that should be addressed in reports of observational studies (<http://www.strobe-statement.org>) (Reprinted with permission)

	Item number	Recommendation
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants (b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/measurement	8 ^a	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses

(continued)

Table 3.12 (continued)

Results	Item number	Recommendation
Participants	13 ^a	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram
Descriptive data	14 ^a	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) Cohort study—Summarize follow-up time (e.g., average and total amount) Cohort study—Report numbers of outcome events or summary measures over time Case—control study—Report numbers in each exposure category, or summary measures of exposure Cross-sectional study—Report numbers of outcome events or summary measures
Outcome data	15 ^a	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarize key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other Information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

^aGive such information separately for cases and controls in case-control studies, and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies. Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of *PLoS Medicine* at <http://www.plosmedicine.org/>, *Annals of Internal Medicine* at <http://www.annals.org/>, and *Epidemiology* at <http://www.epidem.com/>). Separate versions of the checklist for cohort, case-control, and cross-sectional studies are available on the STROBE Web site at <http://www.strobe-statement.org/>

Table 3.13 Randomised controlled trials (RCTs) compared with nonrandomised controlled trials (nonRCTs) of the same intervention (Kunz and Oxman 1998) (Reprinted with permission from BMJ Publishing Group Ltd.)

Study	Sample (search strategy)	Comparison	Results	Direction of bias
Chalmers et al. (1977)	32 controlled studies of anticoagulation in acute myocardial infarction (systematic)	RCTs with CCTs and HCTs on case fatality rate, rate of thromboembolism, and haemorrhages	Relative risk reduction for mortality overestimated by 35% in HCTs and 6% in CCTs compared with RCTs. Case fatality rate highest in HCTs (38.3%) compared with RCTs (19.6%) and CCTs (29.2%). Similar pattern for thromboembolism	Overestimation of effect
Sacks et al. (1982)	Sample of 50 RCTs and 56 HCTs, assessing 6 interventions (treatment of oesophageal varices, coronary artery surgery, anticoagulation in myocardial infarction, chemotherapy for colon cancer and melanoma, and diethylstilboestrol for recurrent miscarriage) (at hand)	RCTs with HCTs on frequency of detecting statistically significant results ($P<0.05$) of primary outcome and reduction of mortality	20% of the RCTs found a statistically significant benefit from the new treatment compared with 79% of the HCTs. Relative risk reduction of mortality in HCTs v RCTs was 0.49/0.27 (1.8) for cirrhosis, 0.68/0.26 (2.6) for coronary artery surgery at 3 years, 0.49/0.22 (2.2) for anticoagulation in myocardial infarction, and 0.67/–0.02 for diethylstilboestrol in recurrent miscarriage. Outcomes in treatment groups were similar in both designs, but outcomes in control groups were worse among historical controls	Overestimation of effect
Diehl and Perry (1986)	19 RCTs and 17 HCTs for 6 types of cancer (breast, colon, stomach, lung cancer, melanoma, soft tissue sarcoma) (reference lists of two textbooks)	Matching of randomised and historical controls for disease, stage, and follow up, and comparison on survival and relapse free survival	18 of 43 matched control groups (42%) varied by >10% (absolute difference in either outcome), 9 (21%) by >20%, and 2 (5%) by >30%. Survival or relapse free survival was better in RCTs compared with HCTs in 17/18 matches	Overestimation of effect
Reimold et al. (1992)	6 RCTs and 6 CCTs of chinidine in atrial fibrillation (systematic)	RCTs and CCTs on maintenance of sinus rhythm 3, 6, and 12 months after cardioversion	At 3 months, beneficial effect of maintaining sinus rhythm with chinidine was 54% less in nonRCTs compared with RCTs, and was 76% less at 12 months	Underestimation of effect
Recurrent Miscarriage Immunotherapy Trialists Group (1994)	9 RCTs and 6 CCTs (with self selected treatment) of allogenic leucocyte immunotherapy for recurrent miscarriage (systematic)	RCTs and CCTs on live birth rate	Beneficial effect of immunotherapy on birth rate among pregnant women was 9% larger in CCTs compared with RCTs, but was 63% lower in CCTs when all women were considered	Underestimation of effect when all women considered, similar effect for pregnant women

(continued)

Table 3.13 (continued)

Study	Sample (search strategy)	Comparison	Results	Direction of bias
Watson et al. (1994)	4 RCTs and 6 CCTs/HCTs of oil soluble contrast media during hysterosalpingography in infertile couples (systematic)	RCTs and CCTs/HCTs on pregnancy rate	RCTs and CCTs/HCTs detected similar increases in pregnancy rates: odds ratio for RCTs 1.92 (95% CI, 1.33–2.68) and for CCTs/HCTs 1.92 (1.55–2.38)	Similar effect
Pyörälä et al. (1995)	11 RCTs and 22 (not further specified) nonRCTs on hormonal therapy in cryptorchidism (systematic)	RCTs and nonRCTs on the descent of testes after therapy with luteinising hormone releasing hormone or human chorionic gonadotrophin	Success rate of descent of testes after therapy with luteinising hormone releasing hormone was 2.3 times larger in nonRCTs than in RCTs and 1.7 times larger after therapy with human chorionic gonadotrophin	Overestimation of effect
Carroll et al. (1996)	17 RCTs and 19 nonRCTs (including HCTs or trials with inadequate randomisation procedures) on transcutaneous electrical nerve stimulation (systematic)	RCTs and nonRCTs on control of postoperative pain	Transcutaneous electrical nerve stimulation judged ineffective at improving postoperative pain in 85% of RCTs, while 89% of nonRCTs concluded that it did improve postoperative pain	Overestimation of effect

CCT concurrently controlled trial, HCT historically controlled trial

can be argued that estimates of effect might be larger in randomized trials if the care provided in the context of trials is better than that in routine practice, assuming this is the case for the treatment group and not the control group. Similarly, strict eligibility criteria might select people with a higher capacity to benefit from a treatment, resulting in larger estimates of effect in randomized trials than nonrandomized trials with less strict eligibility criteria. If, for some reason, patients with a poor prognosis were more likely to be allocated to the treatment group in nonrandomized trials, then this would also result in larger estimates of effect in randomized trials. Conversely, if patients with a poor prognosis were more likely to be allocated to the control group in nonrandomized trials, as often seems to be the case based on the results of this review, this would result in larger estimates of effect in the nonrandomized trials. It was concluded that failure to use adequately concealed random allocation can distort the apparent effects of care in either direction, causing the effects to seem either larger or smaller than they really are (Kunz and Oxman 1998).

Two subsequent Cochrane systematic reviews assessed the effects of randomization and concealment of allocation on the results of health-care studies (Odgaard-Jensen et al. 2011; Kunz et al. 2007). The results of randomized and nonrandomized studies sometimes differed. In some instances, non-randomized studies yielded larger estimates of effect, and in other instances, randomized trials yielded larger estimates of effect. The results of controlled trials with adequate and inadequate/unclear concealment of allocation sometimes differed. When differences occurred, most often, trials with inadequate or unclear allocation concealment yielded larger estimates of effects relative to controlled trials with adequate allocation concealment. However, it is not generally possible to predict the magnitude, or even the direction, of possible selection biases and consequent distortions of treatment effects from studies with non-random allocation or controlled trials with inadequate or unclear allocation concealment (Odgaard-Jensen et al. 2011).

Transparent reporting of intervention trials is important for research synthesis as it facilitates a more accurate critical appraisal and interpretation of the findings. Guidelines for transparent reporting are needed to highlight the key elements necessary for evaluating the validity and strength of effects from intervention trials. The Consolidated Standards of Reporting Trials (CONSORT) Statement has been developed and widely adopted to help guide summarizations of randomized controlled trials. However, behavioural and social interventions, key to such areas as human immunodeficiency virus (HIV) prevention, do not always lend themselves to the format of randomized controlled trials. A parallel set of guidelines for non-randomized controlled trials were developed: **Transparent Reporting of Evaluations with Nonrandomized Designs (TREND Statement)** published in the March 2004 issue of the *American Journal of Public Health*. The TREND Statement provides background on the needs and the process of developing standardized reporting for non-randomized controlled trials and includes a 22-item checklist of elements critical for describing the methods and results of these trials. The TREND Statement should be considered an evolving document as it will be periodically revised as needed. The TREND group welcomes comments and feedback, and more information can be found at www.TREND-statement.org (Vlahov 2004). The CONSORT flow diagram also should be completed.

Examples

For example, Gomes-Filho et al. (2010) described the preliminary results from an alternative methodological resource: the use of a non-randomized design with multiple controls, with emphasis on the historical control group. Through the study hypothesis, a group of 54 pregnant women who underwent periodontal treatment during pregnancy was investigated. These women were compared with another group of 112 women in the immediate postpartum period who had not had previous therapy. It was ensured that all of these subjects came from the same population of preg-

nant women who were attending one of the public health-care clinics in the municipality of Feira de Santana. It was observed that low-weight births occurred twice as frequently among the untreated women than among the women in the test group, who had undergone periodontal therapy during their pregnancies. In the present study, it was found that the test group and control group II were dissimilar regarding the distribution of some covariables, for example, education level, conjugal situation, number of people in the home, smoking habit, and alcohol consumption. In situations like these, indicating lack of comparability between the groups, appropriate use of analytical methods is required to control for unequally distributed baseline factors. Such factors could have been eliminated through randomization (Gomes-Filho et al. 2010).

This alternative methodology has attracted many criticisms because of the existence of sources of non-equivalence between the test and control groups, dispersed over time, thereby giving rise to violation comparability. Sacks et al. (1982) compared the use of randomized controls (RCTs) and historical controls (HCTs) for clinical trials. Six therapies for which 50 RCTs and 56 HCTs were reported. Forty-four of 56 HCTs (79%) found the therapy better than the control regimen, but only 10 of 50 RCTs (20%) agreed. For each therapy, the treated patients in RCTs and HCTs of the same therapy were largely due to differences in outcome for the control groups, with HCT control patients generally doing worse than the RCT control groups. Adjustment of the outcomes of the HCTs for prognostic factors, when possible, did not appreciably change the results. The data suggest that biases in patient selection may irretrievably weight the outcome of HCTs in favour of new therapies. RCTs may miss clinically important benefits because of inadequate attention to sample size. The predictive value of each might be improved by reconsidering the use of P less than 0.05 as the significance level for all types of clinical trials and by the use of confidence intervals around estimates of treatment effects (Sacks et al. 1982).

3.4.2 Randomized Controlled Trial: Gold Standard

The randomized controlled trial is the only known way to avoid selection and confounding biases in clinical research. This design approximates the controlled experiment of basic science. It resembles the cohort study in several respects, with the important exception of randomization of participants to exposures (Grimes and Schulz 2002a).

The randomized clinical trial (RCT) is a research design used to compare two or more therapies. Four traits that distinguish a RCT from routine clinical care are as follows: (1) A RCT is *controlled* (all subjects in the RCT are selected based upon stringent predetermined entrance criteria) and (2) is *randomized* (subjects are assigned to treatment groups by chance, independent of their individual characteristics); (3) the final outcome is established through *statistical analysis* (criteria are established for measuring differences among the two groups prior to initiation of the study, and the data are analysed to determine if the preset level of significance is obtained); and (4) most often, a *double-blind technique* is employed (neither the patient nor the examiner collecting data knows which treatment an individual received) (Dennison 1997).

They are the most effective at demonstrating efficacy of a new intervention or treatment. To bring a new pharmaceutical product to market, the Food and Drug Administration (FDA) will require the compelling evidence of efficacy shown in a true experimental research design study (i.e. a prospective, randomized, controlled, blinded clinical trial). On the downside, these types of studies are the most demanding in terms of time, cost, and other resources. Also, these studies are focused and only can look at a narrow research question (Panacek and Thompson 1995).

3.4.3 Clinical Trials Designs

The study design is the general plan for setting up and testing a specific hypothesis or research question. Overall, the design directs the who, what,

Table 3.14 Research design structure (Thompson and Panacek 2006) Reprinted with permission from Elsevier)

Three research variables
• Independent variable (the intervention) • Dependent variable (the outcome) • Extraneous variables (other, potentially confounding factors)
Three design elements
• Manipulation (the ability to influence or direct the independent variable) • Control (the ability to direct or influence important extraneous variables and study measurements) • Randomization (unbiased [random] subject assignment to each group)

Table 3.15 Classification of research designs by degree of scientific rigour

True experimental
• Have all three design elements • Are always prospective • Have high scientific validity
Quasi-experimental
• Only have one or two design elements • Have manipulation or control • Generally lack randomization • Are generally prospective in nature • Are moderate in scientific validity
Non-experimental
• Have one or none of the core design elements • Lack manipulation and randomization • May also lack control • Are generally retrospective • Have the lowest scientific validity

Thompson and Panacek (2006). Reprinted with permission from Elsevier

when, and how of the research project. Some basic definitions should be reviewed as three important variables apply to all designs, while three other elements are important in understanding study designs: manipulation, control, and randomization (Table 3.14) (Thompson and Panacek 2006).

Table 3.15 summarizes study designs that use the classification system ‘degree of scientific rigor’. **True experimental designs** have all three of the primary characteristics (i.e. manipulation, randomization, and control). They can be thought of as being most similar to a highly controlled

laboratory experiment. As such, these designs have the most safeguards against sources of bias and therefore the greatest degree of overall scientific validity. **Quasi-experimental designs** are missing one or two of these elements. They have the element of manipulation or control but rarely randomization. These designs have the next highest level of scientific validity after ‘true experimental’ but have scientific limitations. As a result, their results are more prone to bias and therefore have less validity. **Non-experimental designs** always lack randomization and usually one or both of the other primary characteristics as well. These studies are generally ‘ex post facto’, that is, retrospective-type designs. Because the study elements of interest have already occurred, it is not possible to randomly assign subjects for the purposes of this specific study. As a result, they have the lowest level of scientific validity, and, to variable degrees, the findings are always open to question (Thompson and Panacek 2006).

True experimental research designs are always prospective in nature. A true experiment can effectively argue a proven cause-and-effect relationship. They are the most effective at demonstrating efficacy of a new intervention or treatment. Many true experimental research designs exist, but they have generally similar structures. The most common ones are outlined in Fig. 3.19. In those descriptions, observation (O) can be any measurement or other data collection activity. Intervention (X) is the manipulation of the study intervention such as drugs, new diagnostic studies, and so forth. The control or comparison intervention (C) could be placebo or current standard care, depending on the study (Thompson and Panacek 2006).

Periodontal clinical trials include patients for whom multiple oral sites potentially qualify for treatment with a test intervention. One design option for these trials is to specify the random assignment of only one intervention per patient such that all qualifying sites within the same patient have the same treatment. This option is called a **parallel-group design** and features two or more groups of patients corresponding to their respective interventions, recruited simultaneously at baseline and followed longitudinally.

Fig. 3.19 True experimental designs (Panacek and Thomason 1995. Reprinted with permission from Elsevier)

Two arm R O X O R O O	Three arm R O X₁ O R O X₂ O R O O
Follow up R O X O₁ O₂ R O O₁ O₂	Factorial R O X₁ O R O X₂ O R O X₁ X₂ O R O X₂ X₁ O R O O
Crossover R O X₁ O X₂ O R O X₂ O X₁ O	
R, randomization; O = observation; X, intervention	

The principal advantage with the parallel groups design is that it enables response differences between the groups to be reasonably interpretable as due to the interventions. This is particularly true when baseline comparability of the groups is assured by randomization and avoidance of post-randomization bias is assured by study masking and monitoring. Moreover, this advantage applies whether data analysis includes only qualifying sites or is for all sites within patients or whether it focuses on patients through summary measures for either qualifying sites or all sites (Koch and Paquette 1997).

The most common parallel research designs are outlined in **Fig. 3.19**. The classic **two-arm design** is that which would be used to study an intervention, such as a new drug. The design uses a placebo control and examines the outcome in two different groups of patients, one that receives the new intervention and one that receives the placebo. A **three-arm study** is very similar but compares two different drugs to placebo. **The extended follow-up design** takes the principles of the two-arm design and makes multiple subsequent observations over a longer period of time (e.g. hospital admission rates, length of stay, relapse rates). This design takes longer to perform but can provide important outcome data. **The factorial design** looks at the effect of multiple interventions, both individually and in various combinations. For example, two different drugs can be evaluated for individual effects, as well as their cumulative effects, depending on the order in

which they are given. This design is attractive to examine each intervention, as well as sequences, but can be very resource expensive because the additional study arms require more total study patients (Thompson and Panacek 2006).

The Crossover Design

Unlike the parallel group trial, crossover trials provide each participant with two or more sequential treatments in a random order usually separated by a washout period. Within a trial, each participant is able to act as his or her own control and permits between and within group comparisons. Crossover studies are most appropriate in studies where the effects of the treatment(s) are short-lived and reversible and are best suited to trials related to symptomatic but chronic conditions or diseases. It is generally agreed that the crossover design should not be used when the condition of interest is unstable and may change regardless of interventions (Mills et al. 2009 and references therein).

The crossover design has numerous advantages that investigators may wish to use for early-stage trials. The particular strength of this design is that the interventions under investigation are evaluated within the same patients and so eliminates between-subject variability. Further, this trial design permits opportunities of head-to-head trials, and patients receiving multiple treatments can express preferences for or against particular treatments (Mills et al. 2009 and references therein).

However, even when properly applied, crossover trials may have certain weaknesses. Patients

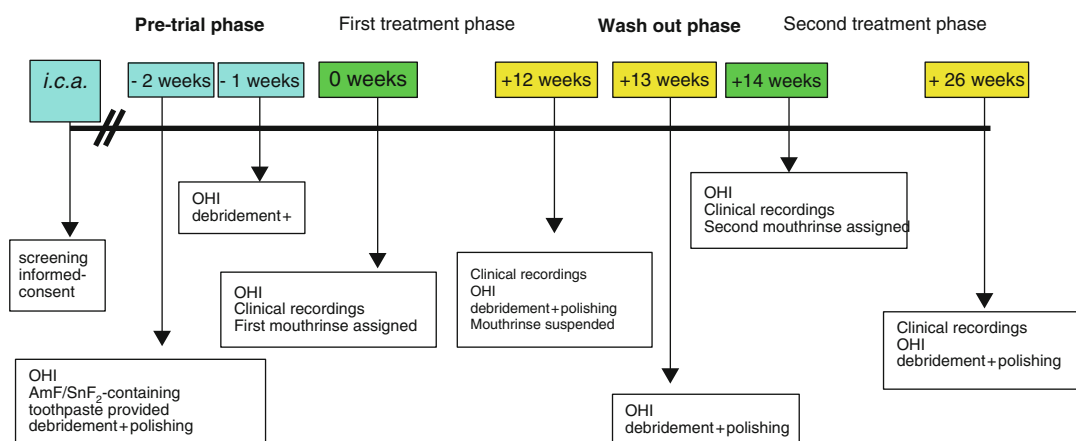


Fig. 3.20 Experimental design and procedures (*i.c.a.* initial contact appointment, *OHI* oral hygiene instructions, *AmF/SnF₂* amine fluoride/stannous fluoride) (Guarnelli

et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

may drop out after the first intervention period and thus not receive a second or third treatment. This makes within-subject comparison impossible and is particularly important if withdrawal is related to side effects. Also, there may be a residual or carry-over of effect of treatments across study periods, which could potentially distort the results obtained during the second treatment or subsequent periods. Thus, the observed treatment effects will depend upon the order in which they were received. Another major potential threat to the validity of the crossover design involves the use of inappropriate statistical analysis. Given that subjects act as their own controls, the analyses could be based on paired data (using an unpaired test), and the within-subject variability in outcomes could be considered in sample size calculations (Mills et al. 2009 and references therein; McCracken et al. 2005).

Parallel-group trials are considered not to be susceptible to **carry-over effects** from previous treatments and to time effects, and so they are currently considered golden standard rather than crossover trials. Just like crossover trials, however, they do have their flaws. For example, in many parallel-group studies, one or more group characteristics are significantly different which may or may not have influenced the results of the study. Parallel-group studies also lack the extra-sensitivity provided by a crossover design which increases with an increasing correlation between

the responses to treatment. Although the first flaw has been widely recognized, little attention has been paid to the second flaw so far (Cleophas and de Vogel 1998). Cleophas and de Vogel (1998) demonstrated that due to this flaw, it is more appropriate to perform a crossover than a parallel-group study for comparing equivalent treatments.

Examples

A recent monocentre, randomized, crossover, double-blind, controlled clinical trial evaluated the efficacy of a mouth rinse containing a combination of AmF/SnF₂ in controlling supra-gingival plaque accumulation and gingival inflammation during a 12-week period in patients affected by generalized aggressive periodontitis (Guarnelli et al. 2004). The study design is summarized in Fig. 3.20. During the pre-trial phase, the patients were provided with oral hygiene instructions (OHI), a medium toothbrush (Elmex Inter X, GABA International AG, Münchenstein, CH), and interdental cleaning devices as needed. An AmF/SnF₂-containing toothpaste (meridol toothpaste, GABA International AG, Münchenstein, CH) was also dispensed. The patients were asked to use the toothpaste three times a day during morning, noon, and evening toothbrushing. At week 0 (baseline), the patients were assigned treatment mouth rinses according to a randomization list. Assignment was performed by a central

Parallel design Randomize N subjects Treatment A Treatment B	Cross-over design Randomize N subjects Treatment A Treatment B x x Treatment B Treatment A
Split-mouth design Randomize k sections Treatment A Treatment B	Split-mouth/cross-over Randomize k sections Treatment A Treatment B x x Treatment B Treatment A

Fig. 3.21 Research designs illustrated for two treatment comparisons (Antczak-Bouckoms et al., 1990. Reprinted with permission from John Wiley & Sons. Inc.)

randomization facility, and examiners were kept unaware of the randomization sequence (allocation concealment). One-half of the patients were either prescribed an AmF/SnF₂-containing mouth rinse (test mouth rinse; meridol mouth rinse, GABA International AG, Münchenstein CH) or a non-AmF/SnF₂-containing mouth rinse (control mouth rinse). Both test and control mouth rinses were prescribed 10 ml twice daily, after morning and evening toothbrushing, for 12 weeks (from week 0 to week +12). OHI, including use of AmF/SnF₂-containing toothpaste, were reinforced. At completion of the first 12-week treatment phase, a 2-week washout phase elapsed (from week +12 to week +14). During the washout phase, the patients reversed to the oral hygiene regimen followed during the pre-trial phase. At week +12 and +13, the patients received OHI, polishing, and ultrasonic debridement as needed for plaque/calculus/stain elimination and gingival inflammation resolution. At week +14 (baseline), the patients received the alternative mouth rinse for 12 weeks according to a crossover design. The patients who had received the test mouth rinse during the first treatment phase received the control mouth rinse during the second treatment phase, and vice versa. The patients were prescribed the AmF/SnF₂-containing toothpaste also throughout this second treatment phase. At week 126, an additional session of polishing and ultrasonic debridement was given. The test and control mouth rinse formulations were identical except for the AmF/SnF₂ content (Guarnelli et al. 2004).

3.4.4 Split-Mouth Trials

Introduced by Ramfjord in periodontal clinical trials (Ramfjord et al. 1968), the split-mouth clinical trial design is a popular design in oral health research (Pihlstrom and Barnett 2010). The split-mouth design is a dental version of an agricultural **split-plot design** where the geographical plots are replaced by regions in the mouth. The split-mouth design is also related to the **crossover design**, where each patient receives more than one treatment in a randomized order. The simplest case is where half of the patients are randomized to the treatment sequence A-B (in the first period, treatment A is administered, and in the second period, treatment B) and the other half to treatment sequence B-A. Thus, one could view the split-mouth design as a kind of crossover design where 'time' is replaced by 'site' in the mouth (Fig. 3.21) (Lesaffre et al. 2009). In periodontology, these designs have been used to evaluate a variety of preventive and therapeutic agents (Antczak-Bouckoms et al. 1990).

The most important assumption underlying the use of this type of design is that the procedure or treatment effects be localized (Antczak-Bouckoms et al. 1990). When there is a leakage of the treatment effect from one site to another site, called a **carry-across effect**, the split-mouth design is seriously handicapped to provide an unbiased estimate of the treatment effect. In a crossover design, this leakage is called the **carry-over effect** and arises because the effect of the

treatment administered first has not completely worn out in the second period. Also, here the leakage is unidirectional, that is, it occurs only from the first to the second period. By increasing the duration between the first and second period in a crossover design, that is, **the washout period**, the carry-over effect can be minimized or eliminated. However, eliminating or controlling a carry-across effect in a split-mouth study is more complicated, if not impossible, because physical barriers cannot be implemented in the mouth. Further, the carry-across effect is typically bidirectional, that is, the treatment administered at site 1 can affect the measurements made at site 2, but equally so the treatment administered at site 2 can affect measurements made at site 1. Finally, while the carry-over effect in a crossover design can be statistically estimated and tested (although with a relatively large uncertainty), it is almost impossible to estimate/test the carry-across effect in a split-mouth study because of the bidirectional effect of the leakage (Lesaffre et al. 2009). **Spillover-effects:** When paired units (teeth, quadrants, arches) are used to compare treatment effects in split-mouth studies, the therapeutic effect of one treatment should not influence the effect of other treatments in the mouth. For example, in studies comparing surgical and non-surgical treatment for periodontal disease, the effect of surgical treatment may be influenced by the fact that microorganisms from sections not receiving surgery may more rapidly re-colonize the surgically treated area. Alternatively, sections receiving non-surgical treatment alone may not provide an unbiased assessment if the patient is unable to effectively remove plaque because of the discomfort following surgery in another section (Antczak-Bouckoms et al. 1990).

Four issues should be considered when contemplating the use of a split-mouth design: bias, recruitment, efficiency, and statistical analysis (Hujoel 1998; Hujoel and Loesche 1990; Hujoel and DeRouen 1992).

1. *Bias*—The split-mouth design may lead to biased treatment efficacy estimates. Bias in split-mouth designs is caused by the confounding of treatment effects with carry-across effects, and the magnitude of this bias cannot be estimated. Randomizing individuals, rather than

split mouths, to treatments eliminates the potential for bias caused by carry-across effects.

2. *Recruitment*—The search for individuals with symmetrical disease patterns can endanger the recruitment process. While it may be easy to find an individual with one intra-osseous lesion of 5 mm, or one three-surface amalgam restoration, it is more difficult to find an individual with two such lesions, especially if symmetrical distribution of the lesions is sought.
3. *Statistical analysis*—The statistical analysis of split-mouth designs is complex. The essential feature of a split mouth is that treatment responses within an individual are correlated. When more than two treatments are investigated, the variance–covariance matrix must satisfy the circularity assumption to allow the usual analysis of variance, and a three-step testing strategy is typically recommended. In longitudinal split-mouth trials, both the correlation between split mouths within a patient and the correlation within the split mouth over time needs to be modelled in the analysis.
4. *Efficiency*—The random allocation of within-patient experimental units to treatments has the potential to eliminate error (*not* bias) from the treatment comparison. The fraction of error eliminated is directly related to the correlation of treatment responses within patients. This correlation is called the **within-patient correlation coefficient**. If only a small fraction of the error term can be eliminated ($r > 0$), the split-mouth design becomes inefficient; even *more* treatment replications may then be required.

Split-mouth trials may also be unknown to statisticians not familiar with oral health research, and it has been reported that many publications do not use the appropriate statistical methods for this design. Because split-mouth studies generate paired responses, both the descriptive statistics as well as the statistical tests should reflect this paired nature (Lesaffre et al. 2007; Hujoel and Moulton 1988).

The analysis of split-mouth and parallel-group studies is not the same. As a result, if a meta-analysis pools both types of trials without considering the differences, the result might be unreliable. In particular, the CI will be incorrect, possibly leading both to an inappropriate conclusion on clinical importance (and statistical significance) and a

Table 3.16 Selection^a of the extended CONSORT guidelines (preliminary version) to clustered data (Lesaffre et al. 2007) (Reprinted with permission from John Wiley & Sons. Inc.)

	Aspect	Guideline	F	D
2	Introduction and background	Reason; rationale for paired design	5	2
6	Outcomes	Clearly defined primary and secondary outcome measures	10	6
7	Sample size	Sample size determination taking into account the correlation between units (range of possible correlations) and treatment crossover effect (range of possible treatment crossover effects)?	1	4
8	Randomization: sequence generation	Methods used to allocate units within a single individual (e.g. how was the first unit to be randomized decided)	14	4
9	Randomization: allocation concealment	Method used to implement the random allocation sequence (e.g. numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned	9	3
11	Blinding (masking)	Whether or not participants, those administering the interventions and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated	15	7
12	Statistical methods	Use of statistical methods appropriate for a paired design	19	7
13	Participant flow	Number of patients and number of units at each stage	7	3
15	Baseline data	Baseline demographic and clinical characteristics of patients (one column); baseline characteristics of units (two columns)	5	4
16	Numbers analysed	Number of randomized units in each group included in each analysis	15	12
17	Outcomes and estimation	Observed correlation in outcomes between sites within individuals for primary outcome and important secondary outcomes	1	3
18	Ancillary analyses	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory	3	8
20	Interpretation	Potential correlation and treatment crossover effect noted	5	6

The column *F* reports the frequency of papers (of 34) that satisfied the guideline to a sufficient extent. The column *D* reports the number of papers that were scored differently by the four scorers

^aThe guidelines were selected because of their statistical nature

distortion of the impact of clinical heterogeneity. Two reviews published on the Cochrane Library provide evidence to support such assertions. Both examined interventions for treating periodontal conditions, guided tissue regeneration (GTR), and enamel matrix proteins (Emdogain) (Needleman et al. 2006; Esposito et al. 2005) and suggested that estimates from split-mouth and parallel-group trials might not be the same (Lesaffre et al. 2009).

Moreover, in the context of superiority trials, a major concern was that the carry-across effect will imply a downward biased effect on the difference in treatment effects. Thus, a carry-across effect in a superiority trial will have a conservative effect on the outcome of the trial. However, in an equivalence or a non-inferiority trial, the effect of a carry-across effect is anti-conservative. Hence, for such studies, even more care must be exercised before embarking on the split-mouth design (Lesaffre et al. 2009). All of these issues mandate that oral

health clinical researchers and statisticians alike carefully consider the implications and limitations of using the split-mouth design in oral health clinical trials (Pihlstrom and Barnett 2010).

In 1988, Hujoel and Moulton reported that only 5 of 22 periodontal split-mouth studies used an appropriate statistical analysis. Either the statistical method was not reported or the authors used an unpaired instead of a paired test. In 2007, Lesaffre et al. evaluated the methodology of split-mouth studies in periodontology after nearly two decades since the last assessment and to extend this across a wide range of clinical research areas in oral health using a preliminary version of the extended CONSORT guidelines for cluster-randomized trials. The results showed that many papers lack essential qualities of good reporting, for example, 5 of 34 papers gave the rationale for choosing a split-mouth design, 19 of 34 (56%) used appropriate analytical statisti-

cal methods, and only 1 of 34 presented an appropriate sample size calculation. Of the five studies that used survival analysis, none of them used a paired approach (**Table 3.16**). In conclusion, despite some progress in analysing and reporting split-mouth studies, there remains a substantial need for improvement.

Examples

Xie et al. (2011) has compared two often used rat models for experimental tooth movement, namely, the elastic band and the coil spring model. They were compared with a silk-ligature-induced periodontitis model. Thirty six, 6-week-old rats were used. In a split-mouth design, four experimental conditions were allocated by permutation: (1) insertion of an orthodontic elastic band between the maxillary first and the second molar; (2) placement of a silk ligature around the cervix of the upper second molar; (3) a 10 cN NiTi coil spring for mesial movement of the three maxillary molars as one block, and (4) control (Fig. 3.22). After 1, 3, and 5 days, rats were killed, and immunohistochemical staining for ED1, cathepsin K, and MMP-9 was performed. A time-dependent increase in the inflammatory infiltration of the interdental papilla was found in the elastic band and the ligature model, but not in the spring model. The disruption of the epithelium of the interdental papilla and the transseptal fibres was less severe in the spring group than in the other

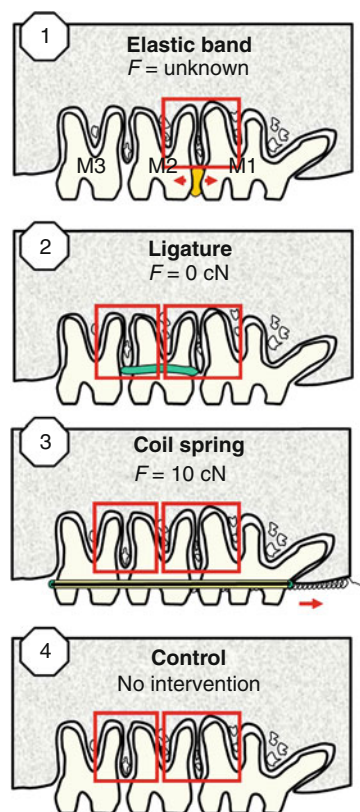


Fig. 3.22 Schematic drawing of the four experimental conditions. 1 – elastic band tooth movement model; 2 – silk ligature periodontitis model; 3 – coil spring tooth movement model; and 4 – control. M1, M2, M3 first, second, and third maxillary molar. F force level. The rectangles mark the observed areas (Xie et al. 2011. Reprinted with permission from Elsevier)

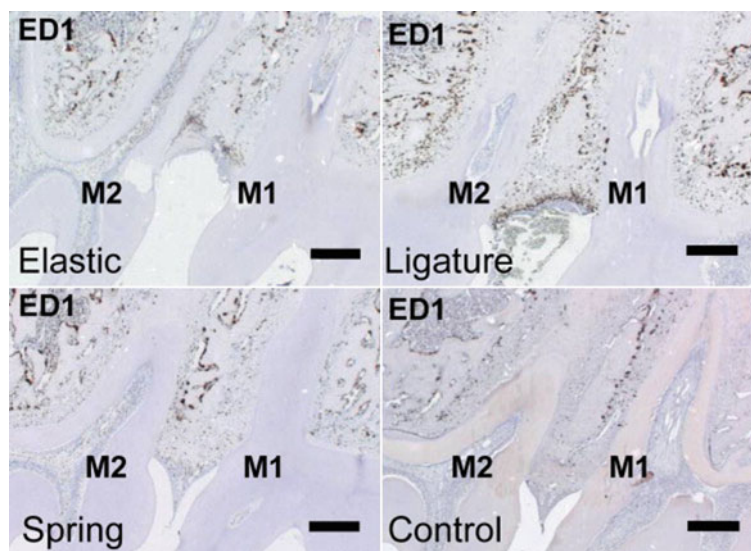


Fig. 3.23 ED1 staining of the interdental region between the first (M1) and second (M2) maxillary molar after 1 day for the four experimental conditions. Bar = 250 μm (Xie et al. 2011. Reprinted with permission from Elsevier)

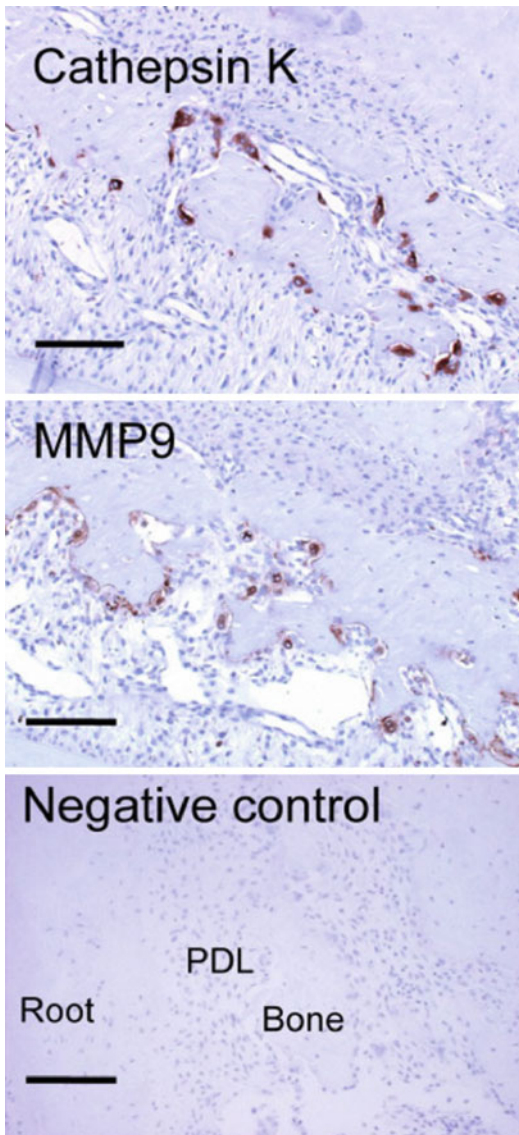


Fig. 3.24 Immunohistochemical staining pattern of cathepsin K, MMP9, and negative control. Bar=100 μ m (Xie et al. 2011). Reprinted with permission from Elsevier)

two experimental groups. All three experimental groups showed a time-dependent resorption of the interdental bony septum. However, this started faster and was more severe in the elastic band and the ligature than in the spring group (Figs. 3.23 and 3.24). The numbers of cathepsin K+ and MMP9+ cells were higher in the elastic band and the ligature group than in the spring group (Xie et al. 2011).

Another split-mouth, randomized, clinical trial aimed to evaluate the efficacy of erbium-doped yttrium–aluminium–garnet (Er:YAG) laser application in non-surgical periodontal treatment. A total of 27 patients underwent four modalities of non-surgical therapy: supra-gingival debridement; scaling and root planing (SRP)1Er:YAG laser; Er:YAG laser; and SRP. Each strategy was randomly assigned and performed in one of the four quadrants (Fig. 3.25, flow chart). Allocation concealment was performed by opaque, sealed envelopes, sequentially numbered. The central registrar generated the allocation sequence by means of a computer-generated random permuted block (patient) and instructed a different subject (investigator) to assign a sealed envelope containing the treatments of each quadrant. The randomization opaque envelope was opened immediately before the beginning of the treatment phase II. Clinical outcomes were evaluated at 3 and 6 months. Six months after therapy, Er:YAG laser showed no statistical difference in clinical attachment gain with respect to supra-gingival scaling [0.15 mm (95% CI –0.16; 0.46)], while SRP showed a greater attachment gain than the supra-gingival scaling [0.37 mm (95% CI 0.05; 0.68)]. No difference resulted between Er:YAG laser+SRP and SRP alone [0.05 mm (95% CI –0.25; 0.36)] (Rotundo et al. 2010).

3.5 Small Clinical Trials

With the exception of large, population-based observational studies that are usually conducted by government agencies, such as the US National Health and Nutrition Examination Survey (NHANES), Australian Research Center for Population Oral Health (ARCPOH), or the United Kingdom Office of National Statistics, most observational studies of oral diseases are relatively small and may be subject to selection bias. The number of participants in a clinical trial should always be large enough to answer the question being posed. In this regard, small-sample clinical trials must be viewed with caution because they may have inadequate statistical

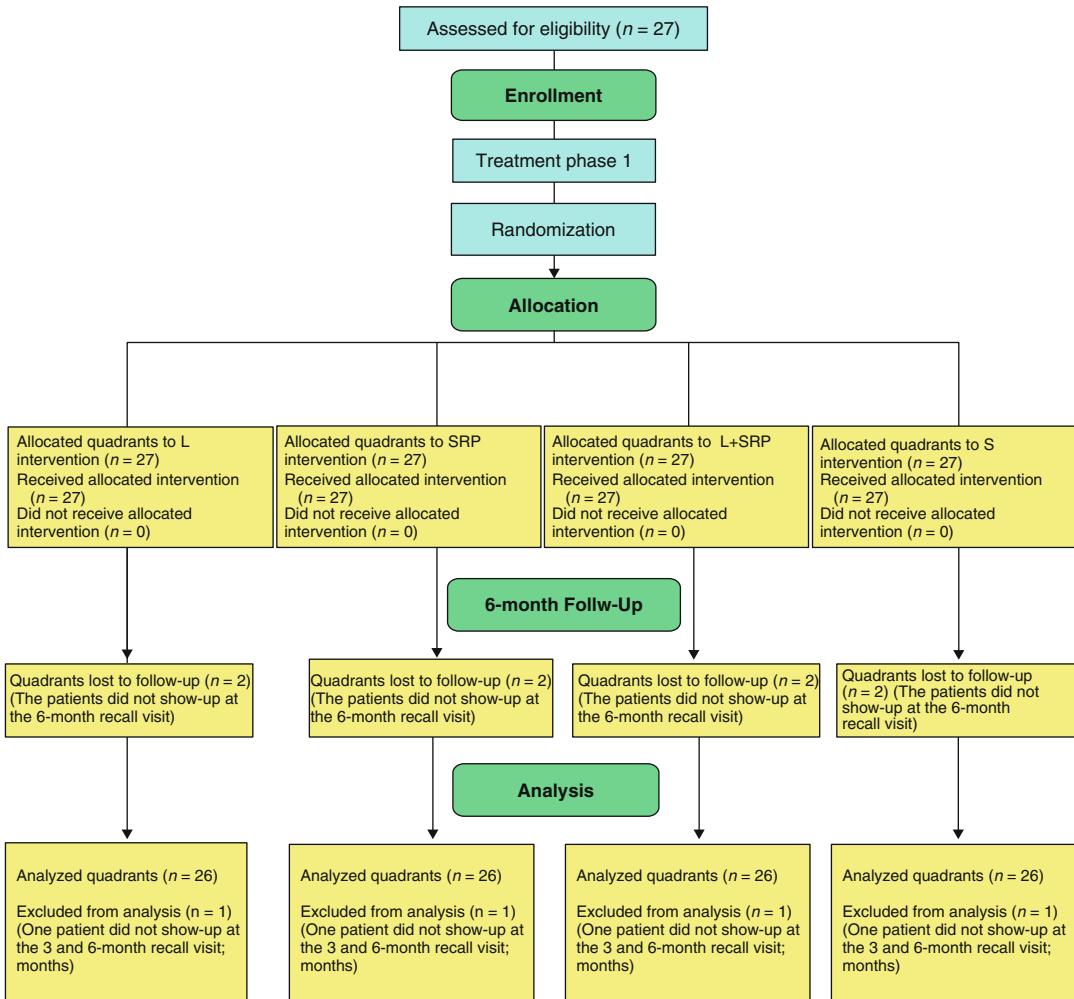


Fig. 3.25 The CONSORT flow-chart of a randomized split-mouth evaluating the adjunctive benefit of Er:YAG laser in non-surgical periodontal treatment (Rotundo et al. 2010. Reprinted with permission from John Wiley & Sons. Inc.)

power to reject the null hypothesis of no difference between the comparative arms of a trial (high false-negative rate). They may also be subject to selection bias or not be generalizable to the overall population because of individual biological variation or variability in methods of health-care delivery (Evans and Ildstad 2001; Pihlstrom and Barnett 2010). In contrast, large clinical trials (hundreds or thousands of participants) may be more likely to be representative of the population as a whole, and their results are more likely to be viewed as generalizable (Pihlstrom and Barnett 2010).

3.6 Sampling Methods

Ensuring that the sample accurately represents the larger population is more important than the size of the sample in quantitative research. The representativeness of the sample in terms of the larger population has a direct impact on the external validity of the conclusions of the research. The first step in sampling then is to identify the **specific characteristics of the target population**, that is, the population to which you wish to generalize the results. The next major consideration in the sampling process is to define the

Table 3.17 Sampling methods in quantitative research (Endacott and Botti 2007) (Reprinted with permission from Elsevier)

Random sampling	
Simple random	All elements in a sampling frame have an equal chance of selection. Selection is through the use of random numbers or similar method
Systematic	Samples are drawn by beginning at a random point in the sampling frame and then selecting each nth element, e.g. Every 20th event or individual. Its advantage is its simplicity
Stratified	Participants are grouped according to strata that are important in a study such as age, sex or diagnosis. Equal numbers of participants are randomly selected from each group
Cluster	The population is divided into smaller sub-sets or clusters. These clusters are randomly selected and all or a random selection of individuals within these clusters are selected. Commonly used in large national surveys to enable random sampling in diverse populations
Non-random sampling	
Convenience	The use of the most readily accessible individuals or units in a study. Used in exploratory research to obtain an estimate of a particular element of interest
Consecutive	A version of convenience sampling where every available individual or event within an accessible population is chosen. The best choice of non-random sampling
Snowball	Individuals recruited for a study refer other potential participants. Useful when participants maybe known within networks rather than generally or when a desired study characteristic is rare
Quota	This method ensures that specific individuals or units are included equally in a convenience sample, for example according to age group or sex

inclusion and exclusion criteria of the accessible population. Inclusion criteria are generally based on the research question and the research plan; they are applied to enable selection of homogenous samples and improve the feasibility of conducting a study. For example, it may be considered that treatment of intrabony defects with enamel matrix proteins, guided tissue regeneration, or open flap debridement may be better studied in patients with (1) presence of one intrabony defect of a probing depth of at least 6 mm, (2) no systemic diseases that could interfere with periodontal healing, (3) no use of antibiotics the least 6 months prior to treatment, and (4) good level of oral hygiene (Sculean et al. 2004). Exclusion criteria are generally applied to exclude unique characteristics that may confound the results or to deal with ethical considerations relating to the research. For example, it may be necessary to exclude patients who have additional medical problems that may affect their outcomes or to exclude patients who have cognitive impairment that impairs their ability to provide informed consent to participate in the study (Endacott and Botti 2007).

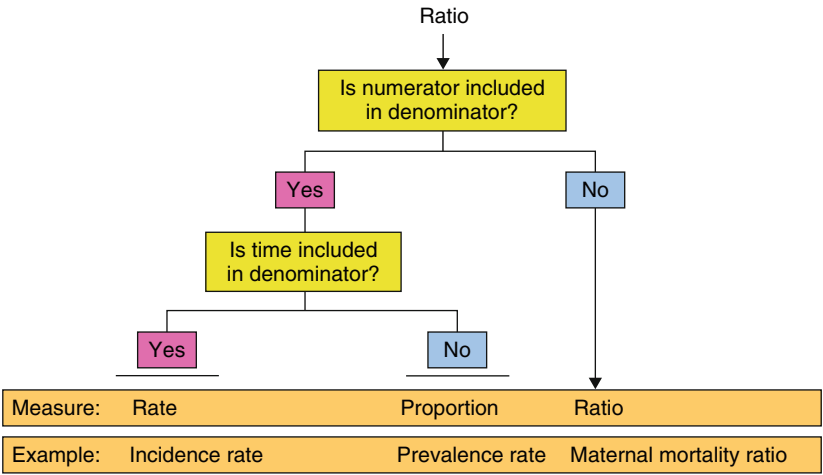
Once the target and accessible population characteristics have been defined, a sample is

selected from the accessible population. There are two main types of samples: random and non-random samples, also referred to as probability and non-probability samples (see **Table 3.17**). **Random or probability sampling** refers to sampling processes that guarantee that each of the potential participants, events, or units under investigation has an equal chance of being selected. Random selection is important for two reasons. First, it reduces the risk of selection bias, and second, a randomly selected sample is a requirement for inferential statistical analyses. In **non-random or non-probability sampling**, the sampling technique is not random; therefore, all members of the population do not have an equal chance of being selected for recruitment into the study. Non-random samples cannot be assumed to fully represent the target population, and consequently, conclusions about the generalizability of results to the target population should be qualified (Endacott and Botti 2007).

Sample size calculation in quantitative research depends on a number of factors (Endacott and Botti 2007). These include:

1. Research design
2. Sampling method

Fig. 3.26 Algorithm for distinguishing rates, proportions, and ratios (Reprinted from Grimes and Schulz 2002a) with permission from Elsevier)



- 3. The degree of precision required
- 4. The variability of the factors being investigated
- 5. The incidence of a particular variable in the population

In descriptive studies, where the intent is to identify the proportion of a particular phenomenon in a sample, sample size calculation is based on the level of precision and confidence required of the results (Endacott and Botti 2007). **In explanatory studies** where hypothesized cause–effect relationships are tested, researchers base sample size calculations on three desired factors: (1) the significance level, (2) power, and (3) effect size. **Qualitative designs** rely on saturation-based sample selection to satisfy the sampling principle of adequacy. Appropriateness of sampling is usually achieved through purposive or theoretical sampling (Endacott and Botti 2007).

3.7 Measurement of Outcomes

3.7.1 Confusing Fractions

Identification and quantification of outcomes is the business of research. However, slippery terminology often complicates matters for investigators and readers alike. **Figure 3.26** presents a simple approach to classification of these common terms (Grimes and Schulz 2002a).

3.7.2 Measures of Association

Relative risk (also termed the risk ratio) is another useful ratio: the frequency of outcome in the exposed group divided by the frequency of outcome in the unexposed. If the frequency of the outcome is the same in both groups, then the ratio is 1.0, indicating no association between exposure and outcome. By contrast, if the outcome is more frequent in those exposed, then the ratio will be greater than 1.0, implying an increased risk associated with exposure. Conversely, if the frequency of disease is less among the exposed, then the relative risk will be less than 1.0, implying a protective effect (Grimes and Schulz 2002a).

Also known as the cross-products ratio or relative odds, **the odds ratio** has different meanings in different settings. In case–control studies, this measure is the usual measure of association. It indicates the odds of exposure among the case group divided by the odds of the exposure among controls. If cases and controls have equal odds of having the exposure, the odds ratio is 1.0, indicating no effect. If the cases have higher odds of exposure than the controls, then the ratio is greater than 1.0, implying an increased risk associated with exposure. Similarly, odds ratios less than 1.0 indicate a protective effect. An odds ratio can also be calculated for cross-sectional, cohort, and randomized controlled studies (**Fig. 3.27**) (Grimes and Schulz 2002a).

	Yes	No
Case	a	b
Control	c	d

$$\text{Odds of exposure among cases} = \frac{a}{b}$$

$$\text{Odds of exposure among controls} = \frac{c}{d}$$

$$\text{Odds ratio} = \frac{a/b}{c/d} = \frac{ad}{cb}$$

Fig. 3.27 The cross-products ratio (odds ratio) in a case-control study (Grimes and Schulz 2002c. Reprinted with permission from Wolters Kluwer Health)

The confidence interval reflects the precision of study results. The interval provides a range of values for a variable, such as a proportion, relative risk, or odds ratio, that has a specified probability of containing the true value for the entire population from which the study sample was taken. Although 95% CIs are the most commonly used, others, such as 90%, are seen (and advocated). The wider the confidence interval, the less precision exists in the result, and vice versa. For relative risks and odds ratios, when the 95% CI does not include 1.0, the difference is significant at the usual 0.05 level. However, use of this feature of confidence intervals as a back-door means of hypothesis testing is inappropriate (Grimes and Schulz 2002a).

3.8 Are the Findings Valid?

When readers encounter research results, they need to consider two basic questions. First, did the study measure what it set out to measure? This characteristic is termed **internal validity**. Assuming the study results have internal validity, the next question to answer is **external validity**: can the results be extrapolated (gener-

alized) to one's patients? (Grimes and Schulz 2002c)

Observational research may have poorer internal validity and better external validity than randomized controlled trials; the opposite is true for randomized controlled trials. Obviously, if a study lacks internal validity, generalizing these invalid results is worthless and possibly misleading (Grimes and Schulz 2002c).

3.9 Assessing the Quality of Published Observational Research

Numerous tools have been proposed for evaluation of methodological quality of observational epidemiological studies. Several comprehensive reviews have generally concluded that there is currently no agreed 'gold standard' appraisal tool, that the majority of tools did not undergo a rigorous development process, and that there are many tools from which to choose (Pladevall-Vila et al. 1996; Juni et al. 2001; Katrak et al. 2004).

Recently, Sanderson et al. (2007) reviewed a total of 86 tools, comprising 41 simple checklists, 12 checklists with additional summary judgments, and 33 scales. The number of items ranged from 3 to 36 (mean 13.7). One-third of tools were designed for single use in a specific review and one-third for critical appraisal. Half of the tools provided development details, although most were proposed for future use in other contexts. Most tools included items for selection methods (92%), measurement of study variables (86%), design-specific sources of bias (86%), control of confounding (78%), and use of statistics (78%); only 4% addressed conflict of interest. The distribution and weighting of domains across tools was variable and inconsistent. The authors concluded that there is a need to agree on critical elements for assessing susceptibility to bias in observational epidemiology and to develop appropriate evaluation tools.

In a recent systematic review investigating the association between overweight or obesity (as defined by the World Health Organization) and

periodontitis (Suvarn et al. 2011), cohort and case-control studies were assessed using the validated Newcastle-Ottawa Quality Assessment Scale (Wells et al. 2009) as recommended by the Cochrane Collaboration guidelines for the assessment of non-randomized studies (Reeves et al. 2009). These tools award stars (*) in three categories for each study based on incorporation of design elements associated with minimizing bias. A validated assessment tool to assess cross-sectional studies was not located. The quality of cross-sectional studies was assessed using questions from the Newcastle-Ottawa Quality Assessment Scale for cohort study design that were deemed applicable to the cross-sectional study design (Wells et al. 2009; Suvarn et al. 2011).

References

- Abel U. Bedeutung von Kasuistiken für die Wirksamkeitsüberprüfung medizinischer Therapiendesigns. In: Böhning MKFH, editor. *Naturheilverfahren und unkonventionelle Medizinische Richtungen*. Berlin: Springer; 1998. p. 1–12.
- Abel U. Erkenntnisgewinn mittels nichtrandomisierter Therapiestudien: gaining understanding through nonrandomized therapeutic studies. *Ellipse*. 1999;15:48–58.
- Akin L, Herford AS, Cicciù M. Oral presentation of disseminated histoplasmosis: a case report and literature review. *J Oral Maxillofac Surg*. 2011;69:535–41.
- Albrecht J, Meves A, Bigby M. Case reports and case series from Lancet had significant impact on medical literature. *J Clin Epidemiol*. 2005;58:1227–32.
- Al-Zahrani MS, Bissada NF, Borawski EA. Obesity and periodontal disease in young, middle-aged, and older adults. *J Periodontol* 2003;74:610–615.
- Al-Zahrani MS, Borawski EA, Bissada NF. Periodontitis and three health-enhancing behaviors: maintaining normal weight, engaging in recommended level of exercise, and consuming a high-quality diet. *J Periodontol* 2005;76:1362–1366.
- Antczak-Bouckoms AA, Tulloch JF, Berkey CS. Split-mouth and cross-over designs in dental research. *J Clin Periodontol*. 1990;17:446–53.
- Bahekar AA, Singh S, Saha S, Molnar J, Arora R. The prevalence and incidence of coronary heart disease is significantly increased in periodontitis: a meta-analysis. *Am Heart J*. 2007;154:830–7.
- Beck J, Garcia R, Heiss G, Vokonas PS, Offenbacher S. Periodontal disease and cardiovascular disease. *J Periodontol*. 1996;67(10 Suppl):1123–37.
- Braegger U. Cost-benefit, cost-effectiveness and cost-utility analyses of periodontitis prevention. *J Clin Periodontol*. 2005;32 Suppl 6:301–13.
- Borges-Yanez SA, Irigoyen-Camacho ME, Maupome G. Risk factors and prevalence of periodontitis in community-dwelling elders in Mexico. *J Clin Periodontol* 2006;33:184–194.
- Braegger U. Cost-benefit, cost-effectiveness and cost-utility analyses of periodontitis prevention. *J Clin Periodontol*. 2005;32 Suppl 6:301–13.
- Briggs JE, et al. Angiographically confirmed coronary heart disease and periodontal disease in middle-aged males. *J Periodontol*. 2006;77:95–102.
- Buhlin K, Gustafsson A, Ahnve S, et al. Oral health in women with coronary heart disease. *J Periodontol*. 2005;76:544–50.
- Carroll D, Tramer M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth*. 1996;77:798–803.
- Chalmers TC, Matta RJ, Smith Jr H, Kunzler AM. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med*. 1977;297:1091–6.
- Chapman A, Munch A, Schermann C, Piacentini CC, Fasolo MT. Obesity and periodontal disease in diabetic pregnant women. *Pesquisa Odontologica Brasileira (Brazilian Oral Res)* 2005;19:83–87.
- Cleophas TJ, de Vogel EM. Crossover studies are a better format for comparing equivalent treatments than parallel-group studies. *Pharm World Sci*. 1998;20:113–7.
- ClinicalTrials.gov. Understanding clinical trials. 2011. <http://clinicaltrials.gov/ct2/info/understand#Q19>. Accessed 30 Apr 2011.
- Dalla Vecchia CF, Susin C, Rosing CK, Oppermann RV, Albandar JM. Overweight and obesity as risk indicators for periodontitis in adults. *J Periodontol* 2005; 76: 1721–1728.
- Deinzer R, Micheelis W, Granrath N, Hoffmann T. More to learn about: periodontitis-related knowledge and its relationship with periodontal health behaviour. *J Clin Periodontol*. 2009;36:756–64.
- Dennison DK. Components of a randomized clinical trial. *J Periodontol Res*. 1997;32:430–8.
- DeStefano F, Anda RF, Kahn HS, Williamson DF, Russell CM. Dental disease and risk of coronary heart disease and mortality [comments]. *Br Med J*. 1993;306: 688–91.
- Diehl LF, Perry DJ. A comparison of randomized concurrent control groups with matched historical control groups: are historical controls valid? *J Clin Oncol*. 1986;4:1114–20.
- Donos N, Glavind L, Karring T, Sculean A. Clinical evaluation of an enamel matrix derivative in the treatment of mandibular degree II furcation involvement: a 36-month case series. *Int J Periodontics Restorative Dent*. 2003;23:507–12.
- Ekuni D, Yamamoto T, Koyama R, Tsuneishi M, Naito K, Tobe K. Relationship between body mass index and

- periodontitis in young Japanese adults. *J Periodontol Res* 2008;43:417–421.
- Endacott R, Botti M. Clinical research 3: sample selection. *Accid Emerg Nurs*. 2007;15:234–8.
- Esposito M, Grusovin M, Coulthard P, Worthington H. Enamel matrix derivative (Emdogain_®) for periodontal tissue regeneration in intrabony defects. *Cochrane Database Syst Rev*. 2005;(4):CD003875. doi:10.1002/14651858.CD003875.pub2. <http://www.ncbi.nlm.nih.gov/pubmed/16235343>
- Evans Jr CH, Ildstad ST. Executive summary. In: *Small clinical trials – issues and challenges*. Washington, DC: Institute of Medicine, National Academy Press; 2001. p. 1–11.
- Geerts SO, Legrand V, Charpentier J, et al. Further evidence of the association between periodontal conditions and coronary artery disease. *J Periodontol*. 2004;75:1274–80.
- Genco R, Chadda S, Grossi S. Periodontal disease is a predictor of cardiovascular disease in a native American population. *J Dent Res*. 1997;76:408.
- Genco RJ, Grossi SG, Ho A, Nishimura F, Murayama Y. A proposed model linking inflammation to obesity, diabetes, and periodontal infections. *J Periodontology* 2005;76(11 Suppl):2075–2084.
- Gjerme PE, Grytten J. Cost-effectiveness of various treatment modalities for adult chronic periodontitis. *Periodontol* 2000. 2009;51:269–75.
- Gomes-Filho IS, Cruz SS, Costa Mda C, Passos JS, Cerqueira EM, Sampaio FP, Pereira EC, Miranda LF. Periodontal therapy and low birth weight: preliminary results from an alternative methodologic strategy. *J Periodontol*. 2010;81(12):1725–33.
- Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet*. 2002a;359(9300):57–61.
- Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002b;359(9302):248–52.
- Grimes DA, Schulz KF. Clinical research in obstetrics and gynecology: a Baedeker for busy clinicians. *Obstet Gynecol Surv*. 2002c;57(9 Suppl 3):S35–53.
- Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. *Lancet*. 2002d;359:341–5.
- Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. *Lancet*. 2002e;359(9301):145–9.
- Grimes DA, Schulz KF. Compared to what? Finding controls for case–control studies. *Lancet*. 2005;365(9468):1429–33.
- Grimes DA. “Case-control” confusion: mislabeled reports in obstetrics and gynecology journals. *Obstet Gynecol*. 2009;114(6):1284–6.
- Guarnelli ME, Zangari F, Manfrini R, Scapoli C, Trombelli L. Evaluation of additional amine fluoride/stannous fluoride-containing mouthrinse during supportive therapy in patients with generalized aggressive periodontitis. A randomized, crossover, double-blind, controlled trial. *J Clin Periodontol*. 2004;31:742–8.
- Han DH, Lim SY, Sun BC, Paek DM, Kim HD. Visceral fat area-defined obesity and periodontitis among Koreans. *J Clin Periodontol* 2010;37:172–179.
- Heasman PA, Vernazza CR, Gaunt FL, Pennington MW. Cost-effectiveness of adjunctive antimicrobials in the treatment of periodontitis. *Periodontol* 2000. 2011;55(1):217–30. doi:10.1111/j.1600-0757.2010.00341.x.
- Hellems MA, Kramer MS, Hayden GF. Case–control confusion. *Ambul Pediatr*. 2006;6:96–9.
- Hennequin-Hoenderdos N, Slot D, Van der Weijden G. Complications of oral and peri-oral piercings: a summary of case reports. *Int J Dent Hyg*. 2011;9:101–9.
- Horton AL, Boggess KA, Moss KL, Beck J, Offenbacher S. Maternal periodontal disease and soluble fms-like tyrosine kinase-1 expression. *J Periodontol*. 2009;80:1506–10.
- Howell T, Ridker P, Ajani U. Periodontal disease and risks of subsequent cardiovascular disease in U.S. male physicians. *J Am Coll Cardiol*. 2001;37:445.
- Hujoel PP. Design and analysis issues in split-mouth clinical trials. *Community Dent Oral Epidemiol*. 1998;26:85–6.
- Hujoel PP, DeRouen TA. Validity issues in split-mouth trials. *J Clin Periodontol*. 1992;19(9 Pt 1):625–7.
- Hujoel PP, Loesche WJ. Efficiency of split-mouth designs. *J Clin Periodontol*. 1990;17:722–8.
- Hujoel PP, Moulton LH. Evaluation of test statistics in split-mouth clinical trials. *J Periodontol Res*. 1988;23:378–80.
- Hujoel PP, Drangsholt M, Spiekerman C, DeRouen TA. Periodontal disease and coronary heart disease risk. *JAMA*. 2000;284:1406–10.
- Janket SJ, Baird AE, Chuang SK, Jones JA. Meta-analysis of periodontal disease and risk of coronary heart disease and stroke. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2003;95:559–69.
- Jenicek M. *Clinical case reporting in evidence-based medicine*. London: Arnold; 2001.
- Joshiyura KJ, Rimm EB, Douglass CW, Trichopoulos D, Ascherio A, Willett WC. Poor oral health and coronary heart disease. *J Dent Res*. 1996;75:1631–6.
- Johansson I, Tidehag P, Lundberg V, Hallmans G. Dental status, diet and cardiovascular risk factors in middle-aged people in northern Sweden. *Community Dent Oral Epidemiol* 1994;22:431–436.
- Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42–6.
- Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol*. 2004;4:22.
- Khoct A, Viera-Negron YE, Ameri A, Abdelsayed R. Periodontitis associated with Chediak-Higashi syndrome in a young African American male. *J Int Acad Periodontol*. 2010;12:49–55.
- Khader YS, Bawadi HA, Haroun TF, Alomari M, Tayyem RF. The association between periodontal disease and obesity among adults in Jordan. *J Clin Periodontol* 2009;36:18–24.
- Kim TS, Holle R, Hausmann E, Eickholz P. Long-term results of guided tissue regeneration therapy with non-resorbable and bioabsorbable barriers. II. A case series of infrabony defects. *J Periodontol*. 2002;73:450–9.
- Koch GG, Paquette DW. Design principles and statistical considerations in periodontal clinical trials. *Ann Periodontol*. 1997;2:42–63.

- Kongstad J, Hvidtfeldt UA, Gronbaek M, Stoltze K, Holmstrup P. The relationship between body mass index and periodontitis in the Copenhagen City Heart Study. *J Periodontol* 2009;80:1246–1253.
- Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomized and non-randomized clinical trials. *BMJ*. 1998;317:1185–90.
- Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev*. 2007;(2):MR000012, <http://www.ncbi.nlm.nih.gov/pubmed/17443633>
- Kushiyama M, Shimazaki Y, Yamashita Y. Relationship between metabolic syndrome and periodontal disease in Japanese adults. *J Periodontol* 2009;80:1610–1615.
- Leiff S, Boggess KA, Murtha AP, et al. The oral conditions and pregnancy study: periodontal status of a cohort of women. *J Periodontol*. 2004;75:116–26.
- Lesaffre E, GarciaZattera MJ, Redmond C, Huber H, Needleman I, ISCB Subcommittee on Dentistry. Reported methodological quality of split-mouth studies. *J Clin Periodontol*. 2007;34:756–61.
- Lesaffre E, Philstrom B, Needleman I, Worthington H. The design and analysis of split-mouth studies: what statisticians and clinicians should know. *Stat Med*. 2009;28:3470–82.
- Levin L, Ofec R, Grossmann Y, Anner R. Periodontal disease as a risk for dental implant failure over time: a long-term historical cohort study. *J Clin Periodontol*. 2011;38:732–7.
- Luepker RV. Observational studies in clinical research. *J Lab Clin Med*. 2005;146:9–12.
- Malthaner SC, et al. Investigation of the association between angiographically defined coronary artery disease and periodontal disease. *J Periodontol*. 2002;73:1169–76.
- Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. *Emerg Med J*. 2003;20:54–60.
- Mattila KJ, Nieminen M, Valtonen V, et al. Association between dental health and acute myocardial infarction. *Br Med J*. 1989a;298:779–81.
- Mattila KJ, Rasi V, Nieminen M, et al. von Willebrand factor antigen and dental infections. *Thomb Res*. 1989b;56:325–9.
- Mattila KJ, Valtonen VV, Nieminen M, Huttunen JK. Dental infection and the risk of new coronary events: prospective study of patients with documented coronary artery disease. *Clin Infect Dis*. 1995;20:588–92.
- Mayo NE, Goldberg MS. When is a case-control study not a case-control study? *J Rehabil Med*. 2009;41:209–16.
- McCracken GI, Steen N, Preshaw PM, Heasman L, Stacey F, Heasman PA. The crossover design to evaluate the efficacy of plaque removal in tooth-brushing studies. *J Clin Periodontol*. 2005;32:1157–62.
- McIntosh CL, Kolhatkar S, Winkler JR, Ojha J, Bhola M. An unusual case of generalized severe gingival enlargement during pregnancy. *Gen Dent*. 2010;58:e272–8.
- Mihailovic A, Bell AM, Urbach DR. Users' guide to the surgical literature: case-control studies in surgical journals. *Can J Surg*. 2005;48:148–51.
- Mills EJ, Chan AW, Wu P, Vail A, Guyatt GH, Altman DG. Design, analysis, and presentation of crossover trials. *Trials*. 2009;10:27.
- Morabia A. Case-control studies in clinical research: mechanism and prevention of selection bias. *Prev Med*. 1997;26:674–7.
- Morrison HI, Ellison LF, Taylor GW. Periodontal disease and risk of fatal coronary heart and cerebrovascular diseases. *J Cardiovasc Risk*. 1999;6:7–11.
- Moses LE. The series of consecutive cases as a device for assessing outcomes of intervention. *N Engl J Med*. 1984;311:705–10.
- Needleman I, Worthington H, Giedrys-Leeper E, Tucker R. Guided tissue regeneration for periodontal infra-bony defects. *Cochrane Database Syst Rev*. 2006;(2):CD001724. doi:10.1002/14651858.CD001724.pub2, <http://www.ncbi.nlm.nih.gov/pubmed/17443633>
- Newton JT, Bower EJ, Williams AC. Research in primary dental care. Part 2: Developing a research question. *Br Dent J*. 2004;196:605–8.
- Nishida N, Tanaka M, Hayashi N, Nagata H, Takeshita T, Nakayama K, Morimoto K, Shizukuishi S. Determination of smoking and obesity as periodontitis risks using the classification and regression tree method. *J Periodontol* 2005;76:923–928.
- Nucci Da Silva L, Gun C, Simone J. Strong Association Between Periodontal Disease And Obesity Among Brazilians. *Atheroscler Suppl* 2009; 10: 1389. Abstract.
- Odgaard-Jensen J, Vist GE, Timmer A, Kunz R, Akl EA, Schünemann H, Briel M, Nordmann AJ, Pregno S, Oxman AD. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev*. 2011;(4):MR000012. <http://www.ncbi.nlm.nih.gov/pubmed/?term=Odgaard-Jensen%20Randomisation%20to%20protect%20against%20selection%20bias%20in%20healthcare%20trials>
- Okuda K, Yamamiya K, Kawase T, Mizuno H, Ueda M, Yoshie H. Treatment of human infrabony periodontal defects by grafting human cultured periosteum sheets combined with platelet-rich plasma and porous hydroxyapatite granules: case series. *J Int Acad Periodontol*. 2009;11:206–13.
- Olson SH, Kelsey JL, Pearson TA, Levin B. Evaluation of random digit dialing as a method of control selection in case-control studies. *Am J Epidemiol*. 1992;135:210–22.
- Östberg AL, Nyholm M, Gullberg B, Rastam L, Lindblad U. Tooth loss and obesity in a defined Swedish population. *Scand J Public Health* 2009;37:427–433.
- Panacek EA. Survey-based research: general principles. *Air Med J*. 2008a;27:14–6.
- Panacek EA. Survey-based research: performing the survey. *Air Med J*. 2008b;27:64–6.
- Panacek EA, Thomason CB. Basics of research (part 3): research study design. *Air Med J*. 1995;14:139–46.
- Pearson TA, Mensah GA, Alexander RW, et al. Markers of inflammation and cardiovascular disease: applica-

- tion to clinical and public health practice: a statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation*. 2003;107:499–511.
- Pihlstrom BL, Barnett ML. Design, operation, and interpretation of clinical trials. *J Dent Res*. 2010;89:759–72.
- Pladevall-Vila M, Delclos GL, Varas C, Guyer H, Bragues-Tarradellas J, Anglada-Arisa A. Controversy of oral contraceptives and risk of rheumatoid arthritis: meta-analysis of conflicting studies and review of conflicting meta-analyses with special emphasis on analysis of heterogeneity. *Am J Epidemiol*. 1996;144:1–14.
- Pradeep AR, Manojkumar ST, Arjun R. Pemphigus vulgaris with significant periodontal findings: a case report. *J Calif Dent Assoc*. 2010;38:343–6.
- Pretzl B, Kim TS, Holle R, Eickholz P. Long-term results of guided tissue regeneration therapy with non-resorbable and bioabsorbable barriers. IV. A case series of infrabony defects after 10 years. *J Periodontol*. 2008;79:1491–9.
- Pyörälä S, Huttunen NP, Uhari M. A review and meta-analysis of hormonal treatment of cryptorchidism. *J Clin Endocrinol Metab*. 1995;80:2795–9.
- Ramfjord SP, Nissle RR, Shick RA, Cooper H Jr. Subgingival curettage versus surgical elimination of periodontal pockets. *J Periodontol*. 1968;39(3):167–75.
- Recurrent Miscarriage Immunotherapy Trialists Group. Worldwide collaborative observational study and meta-analysis on allogenic leukocyte immunotherapy for recurrent spontaneous abortion. *Am J Reprod Immunol*. 1994;32:55–72.
- Reeves B, Deeks J, Higgins J, Wells G; Cochrane Non-Randomised Studies Methods Group. Including non-randomized studies. 2009. Report.
- Reimold SC, Chalmers TC, Berlin JA, Antman EM. Assessment of the efficacy and safety of antiarrhythmic therapy for chronic atrial fibrillation: observations on the role of trial design and implications of drug related mortality. *Am Heart J*. 1992;124:924–32.
- Rocuzzo M, De Angelis N, Bonino L, Aglietta M. Ten-year results of a three-arm prospective cohort study on implants in periodontally compromised patients. Part I: implant loss and radiographic bone loss. *Clin Oral Implants Res*. 2010;21:490–6.
- Rosen PS, Reynolds MA. A retrospective case series comparing the use of demineralized freeze-dried bone allograft and freeze-dried bone allograft combined with enamel matrix derivative for the treatment of advanced osseous lesions. *J Periodontol*. 2002;73:942–9.
- Rotundo R, Nieri M, Cairo F, Franceschi D, Mervelt J, Bonaccini D, Esposito M, Pini-Prato G. Lack of adjunctive benefit of Er:YAG laser in non-surgical periodontal treatment: a randomized split-mouth clinical trial. *J Clin Periodontol*. 2010;37:526–33.
- Sacks H, Chalmers TC, Smith Jr H. Randomized versus historical controls for clinical trials. *Am J Med*. 1982;72:233–40.
- Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*. 2007;36:666–76.
- Saito T, Shimazaki Y, Sakamoto M. Obesity and periodontitis. *New England J Medicine* 1998;339:482–483.
- Saito T, Shimazaki Y, Koga T, Tsuzuki M, Ohshima A. Relationship between upper body obesity and periodontitis. *J Dent Res* 2001;80:1631–1636.
- Saito T, Shimazaki Y, Kiyohara Y, Kato I, Kubo M, Iida M, Yamashita Y. Relationship between obesity, glucose tolerance, and periodontal disease in Japanese women: the Hisayama study. *J Periodontol Res* 2005;40:346–353.
- Saxlin T, Suominen-Taipale L, Leiviska J, Jula A, Knuuttila M, Ylostalo P. Role of serum cytokines tumour necrosis factor alpha and interleukin-6 in the association between body weight and periodontal infection. *J Clin Periodontol* 2009;36:100–105.
- Schröen O, Sahrman P, Roos M, Attin T, Schmidlin PR. A survey on regenerative surgery performed by Swiss specialists in periodontology with special emphasis on the application of enamel matrix derivatives in infrabony defects. *Schweiz Monatsschr Zahnmed*. 2011;121:136–42.
- Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet*. 2002;359:431–4.
- Sculean A, Donos N, Schwarz F, Becker J, Brex M, Arweiler NB. Five-year results following treatment of intrabony defects with enamel matrix proteins and guided tissue regeneration. *J Clin Periodontol*. 2004;31:545–9.
- Shimazaki Y, Saito T, Yonemoto K, Kiyohara Y, Iida M, Yamashita Y. Relationship of metabolic syndrome to periodontal disease in Japanese women: the Hisayama Study. *J Dent Res* 2007;86:271–275.
- Socransky SS, Haffajee AD. Periodontal microbial ecology. *Periodontol* 2000. 2005; 38: 135–187.
- Spahr A, et al. Periodontal infections and coronary heart disease: role of periodontal bacteria and importance of total pathogen burden in the Coronary Event and Periodontal Disease (CORODONT) study. *Arch Intern Med*. 2006;166:554–9.
- Stavropoulos A, Karring T. Five-year results of guided tissue regeneration in combination with deproteinized bovine bone (Bio-Oss) in the treatment of intrabony periodontal defects: a case series report. *Clin Oral Investig*. 2005;9:271–7.
- Suvan J, D'Aiuto F, Moles DR, Petrie A, Donos N. Association between overweight/obesity and periodontitis in adults. A systematic review. *Obes Rev*. 2011;12:e381–404.
- Summers CJ, Oberman A. Association of oral disease with 12 selected variables. I. Periodontal disease. *J Dent Res* 1968;47:457–462.

- Szklo M. Population-based cohort studies. *Epidemiol Rev.* 1998;20:81–90.
- Thompson CB, Panacek EA. Research study designs: experimental and quasi-experimental. *Air Med J.* 2006;25:242–6.
- Thompson CB, Panacek EA. Research study designs: non-experimental. *Air Med J.* 2007;26:18–22.
- Thompson CB, Panacek EA. Research Study Designs: Non-experimental. *Air Medical Journal* 2007;26:18–26:2007 doi:10.1016/j.amj.2006.10.003
- Toomarian L, Hashemi N. Periodontal manifestation of leukocyte adhesion deficiency type I. *Arch Iran Med.* 2010;13:355–9.
- Torrunguang K, Tamsailom S, Rojanasomsith K, Sutdhibhisal S, Nisapakultorn K, Vanichjakvong O, Prapakamol S, Premsirinirund T, Pusiri T, Jaratkulangkoon O, Unkurapinun N, Sritara P. Risk indicators of periodontal disease in older Thai adults. *J Periodontol* 2005;76:558–565.
- Totten VY, Panacek EA, Price D. Basics of research (Part 14). Survey research methodology: designing the survey instrument. *Air Med J.* 1999;18:26–34.
- Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology.* 2007;18:805–35.
- Vlahov D. Transparent Reporting of Evaluations with Nonrandomized Designs (TREND). *J Urban Health.* 2004;81:163–4.
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007; 370(9596):1453–7.
- Watson A, Vandekerckhove P, Lilford R, Vail A, Brosens I, Hughes E. A meta-analysis of the therapeutic role of oil soluble contrast media at hysterosalpingography: a surprising result? *Fertil Steril.* 1994;61:470–7.
- Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. WWW document. 2009. http://www.ohri.ca/programs/clinical_epidemiology/oxford_web.ppt. Department of Epidemiology and Community Medicine, University of Ottawa, Canada. Accessed 26 May 2009.
- Williams AC, Bower EJ, Newton JT. Research in primary dental care part 3: designing your study. *Br Dent J.* 2004;196:669–74.
- Wood N, Johnson RB, Streckfus CF. Comparison of body composition and periodontal disease using nutritional assessment techniques: Third National Health and Nutrition Examination Survey (NHANES III). *J Clin Periodontol* 2003;30:321–327.
- Wu T, Trevisan M, Genco R. Periodontal disease as a risk factor for CVD, CHD and stroke. *Circulation.* 1999;99:1109–25.
- Xie R, Kuijpers-Jagtman AM, Maltha JC. Inflammatory responses in two commonly used rat models for experimental tooth movement: comparison with ligature-induced periodontitis. *Arch Oral Biol.* 2011;56: 159–67.
- Ye L, Liu R, White N, Alon US, Cobb CM. Periodontal status of patients with hypophosphatemic rickets – case series. *J Periodontol.* 2011;82(11):1530–5. Epub 2011 Mar 21.
- Ylöstalo P, Suominen-Taipale L, Reunanen A, Knuuttila M. Association between body weight and periodontal infection. *J Clin Periodontol* 2008;35:297–304.
- Yuen HK, Wolf BJ, Bandyopadhyay D, Magruder KM, Salinas CF, London SD. Oral health knowledge and behavior among adults with diabetes. *Diabetes Res Clin Pract.* 2009;86:239–46.
- Zahedi CS, Miremadi SA, Brunel G, Rompen E, Bernard JP, Benque E. Guided tissue regeneration in human Class II furcation defects using a diphenylphosphorylazide-cross-linked collagen membrane: a consecutive case series. *J Periodontol.* 2003;74:1071–9.

The randomized clinical trial (RCT) is a research design used to compare two or more forms of therapy. Four traits that distinguish an RCT from routine clinical care are (1) an RCT is controlled (all subjects in the RCT are selected based upon stringent predetermined entrance criteria), (2) it is randomized (subjects are assigned to treatment groups by chance, independent of their individual characteristics), (3) the final outcome is established through statistical analysis (criteria are established for measuring differences among the two groups prior to initiation of the study, and the data are analysed to determine of the preset level of significance is obtained), (4) most often, a double-blind technique is employed (Dennison 1997).

RCTs are justifiably recognized as the gold standard in clinical research. Despite this, they have well-recognized disadvantages, including high costs, administrative complexity, prolonged time to completion, and recruitment difficulty (Fisher and Wood 2007).

4.1 Clinical Trial Design

4.1.1 Formulation of Null Hypothesis

The RCT is designed to test for significant differences between two or more therapies. Typically, these different therapies are considered to be equivalent, and statistical evaluation is done to prove or disprove this hypothesis. In statistical nomenclature, the hypothesis being tested is the

null hypothesis. The ethical justification for beginning an RCT requires, at a minimum, that the investigators be able to make an honest statement of no difference—that is a statement that there is no scientifically validated reason to predict that therapy A will be superior to therapy B. Further, there must be no therapy C known to be superior to either A or B unless there is a good cause to reject therapy C, for example, the population of research will consist of either those in whom therapy C has been tried without success, or individuals who are aware of therapy C and for reasons they consider important, have refused it (Dennison 1997).

The ethics of clinical research requires equipoise—a state of genuine uncertainty on the part of the clinical investigator regarding the comparative therapeutic merits of each arm in a trial. Should the investigator discover that one treatment is of superior therapeutic merit, he or she is ethically obliged to offer that treatment. The current understanding of this requirement, which entails that the investigator have no ‘treatment preference’ throughout the course of the trial, presents nearly insuperable obstacles to the ethical commencement or completion of a controlled trial and may also contribute to the termination of trials because of the failure to enrol enough patients (Freedman 1987).

Freedman (1987) suggested that the null hypothesis be replaced with an alternative concept of equipoise, which would be based on present or imminent controversy in the clinical community over the preferred treatment.

Table 4.1 Issues that must be addressed in the planning of a clinical trial (Pihlstrom and Barnett 2010) (Reprinted with permission from Sage Publications)

- The type of trial must be defined (efficacy, effectiveness, superiority, equivalence).
- Clinically meaningful primary and secondary outcomes must be identified and clearly stated in writing.
- Feasibility studies must be conducted.
- Sample size and statistical power must be estimated.
- The study population must be defined.
- Inclusion and exclusion criteria for participant enrollment must be developed.
- Plans for recruitment of participants must be developed.
- Co-investigators and enrollment sites must be identified and recruited.
- An experienced data coordinating center must be identified.
- Plans for data collection, management, and statistical analyses must be developed.
- Trial costs must be estimated as precisely as possible.
- A publication policy should be developed that includes authorship guidelines.
- Plans must be made for developing a manual of procedures.
- Institutional Review Board (IRB) or Institutional Ethics Committee (IEC) approval must be obtained.
- Funding must be obtained.
- A Data and Safety Monitoring Board (DSMB) must be established.
- For United States studies involving a new drug, an Investigational New Drug (IND) application must be filed with the US Food and Drug Administration.

According to this concept of **clinical equipoise**, the requirement is satisfied if there is genuine uncertainty within the expert medical community—not necessarily on the part of the individual investigator—about the preferred treatment.

4.1.2 Protocol Development

Important issues that must be addressed in the planning of a clinical trial are outlined in **Table 4.1**, as described by Pihlstrom and Barnett (2010).

4.1.3 Clinical Endpoints and Biomarkers

In a typical study design for investigating a periodontitis therapy, a study population is selected based on age and **clinical signs of periodontitis** (i.e., probing pocket depth and attachment loss), subjects are randomly assigned to either a control or active treatment group, the deepest site or sites in each quadrant are selected for monitoring, and pocket depth and attachment levels are reassessed at defined intervals. The critical issues associated with this design include imprecision in diagnosis

since periodontal diseases with the same clinical presentation may have different microbial aetiologies; unknown subject disease susceptibility; unknown site disease activity, for example, selected sites may reflect past disease activity and be currently inactive; and limitations of periodontal probing as a measurement method, for example, variability as a result of differences in angulation, force, and position on the tooth. Moreover, because of the potential variation in microbial aetiology and disease activity, only a subset of subjects may respond to specific antimicrobial agents, and group means may not reflect the true efficacy of the test agent. In such a case, the study would require a larger subject population to show statistically significant treatment effects (Fine 2004).

To increase study efficiency, subject entry criteria should include a **microbial component**, so that only subjects with the infectious aetiology appropriate for the antimicrobial being tested are included. In addition, only subjects with active disease should be included. This might be achieved by (1) the development of a quantitative DNA/DNA hybridization method compatible with chairside use, which can be customized to include organisms proven to have an etiologic

Table 4.2 Benefits of randomisation (Schulz and Grimes 2002c) (Reprinted from The Lancet 2002; 359(9305):515–9 with permission from Elsevier)

Proper implementation of a randomisation mechanism affords at least three major advantages:

1. *It eliminates bias in treatment assignment*

Comparisons of different forms of health interventions can be misleading unless investigators take precautions to ensure that their trial comprises unbiased comparison groups relative to prognosis. In controlled trials of prevention or treatment, randomisation produces unbiased comparison groups by avoiding selection and confounding biases. Consequently, comparison groups are not prejudiced by selection of particular patients, whether consciously or not, to receive a specific intervention. The notion of avoiding bias includes eliminating it from decisions on entry of participants to the trial, as well as eliminating bias from the assignment of participants to treatment, once entered. Investigators need to properly register each participant immediately on identification of eligibility for the trial, but without knowledge of the assignment. The reduction of selection and confounding biases underpins the most important strength of randomisation. Randomisation prevails as the best study design for study of small or moderate effects.

2. *It facilitates blinding (masking) of the identity of treatments from investigators, participants, and assessors, including the possible use of a placebo*

Such manoeuvres reduce bias after random assignment, and would be difficult, perhaps even impossible, to implement if investigators assigned treatments by a non-random scheme.

3. *It permits the use of probability theory to express the likelihood that any difference in outcome between treatment groups merely indicates chance*

role in disease; and (2) the identification and validation of specific biomarkers for disease activity in gingival crevicular fluid. For clinical trials of host-response modifiers, **individual susceptibility to periodontal disease**, rather than specific microbial aetiology, is a key consideration, and clinical trial efficiency could be enhanced by the incorporation of a measure of disease susceptibility in subject inclusion criteria, to provide a more homogeneous study population with similar rates of disease progression. Since studies with subjects having a more rapid rate of progression could be shorter and need fewer subjects, this modification could be especially useful in proof-of-principle (Phase II) clinical trials. This might be achieved by identifying genetic markers or more likely, patterns of genetic markers that correlate with increased (or decreased) susceptibility. Diagnostic methods other than periodontal probing might allow for the detection of changes in disease status more precisely and in shorter time periods. More precise and sensitive measures would lead to reduced subject variability, thereby reducing the number of subjects required; allow for detection of changes at earlier time periods, thereby reducing the study duration; and by detecting disease at earlier stages, facilitate the study of preventive interventions. For this to be achieved, **biomarkers to serve as surrogate**

endpoints need to be identified and validated, and new, more sensitive or specific imaging techniques should be investigated for applicability to periodontal clinical trials (Fine 2004).

4.1.4 Methods of Randomization

Assignment to treatments by chance, rather than choice, is the defining feature of randomized controlled trials. This simple, yet powerful, tool ensures that the groups under study are similar in all important respects except for the exposure in question. By balancing baseline characteristics, randomization levels the playing field at the start of a study; selection bias is precluded. By having outcome measures defined in advance and determined in the same way for the groups under study, information bias is avoided or minimized. Potential confounding factors should be balanced between the two groups, assuring comparability at the starting blocks, except for chance imbalance. This is important for known confounding factors but even more critical for those factors that are unsuspected. They, too, will be equally distributed among the groups under study (Table 4.2) (Grimes and Schulz 2002c).

The timing of randomized trials was eloquently discussed in a short letter by Chalmers

Table 4.3 Simple randomisation (Schulz and Grimes (2002c) (Reprinted from The Lancet 2002; 359(9305):515–9 with permission from Elsevier)

An almost infinite number of methods can be used to generate a simple randomisation sequence based on a random-number table. For example, for equal allocation to two groups, predetermine the direction to read the table: up, down, left, right, or diagonal. Then select an arbitrary starting point—i.e., first line, 7th number:
56 99 20 20 52 49 05 78 58 50 62 86 52 11 88
31 60 26 13 69 74 80 71 48 73 72 18 60 58 20
55 59 06 67 02 ...
For equal allocation, an investigator could equate odd and even numbers to interventions A and B, respectively. Therefore, a series of random numbers 05, 78, 58, 50, 62, 86, 52, 11, 88, 31, &c, represent allocation to intervention A, B, B, B, B, B, B, A, B, A, &c. Alternatively, 00–49 could equate to A and 50–99 to B, or numbers 00–09 to A and 10–19 to B, ignoring all numbers greater than 19. Any of a myriad of options suffice, provided the assignment probabilities and the investigator adhere to the predetermined scheme.

several years ago (Chalmers 1968). He argued that the initial studies of new drugs in humans should begin with randomization because at this early point, when there are no data on efficacy and safety, randomization is most ethical. Once treatments become established, it is difficult to conduct trials to understand their effectiveness. Using data from non-randomized studies on the association of treatment and disease outcomes (e.g., anti-infectives for periodontal disease and cardiovascular disease) to inform practice can be fraught with problems. The interventions are likely to have only modest effects on health outcomes, and bias and confounding in non-randomized studies might dominate the observed effects, leading to under- or overestimation of benefit. Also, oral health interventions may be associated with long-term risks, particularly if they involve drugs rather than less benign interventions such as flossing (Neaton and Mugglin 2006).

Randomization consists of two components: firstly, generation of a true random sequence and secondly, concealment of the random sequence from investigators selecting individuals for the trial (**allocation concealment**) (Fenwick et al. 2008).

4.1.4.1 Simple (Unrestricted) Randomization

Simple randomisation is the most basic method of random treatment assignment. This can be thought of as tossing a coin for each trial participant, A being allocated with ‘heads’, B with ‘tails’. However, it is not usually performed using a real coin toss, as issues of concealment, validation, and reproducibility arise. Simple random-

ization is usually achieved using a **sequence of random numbers** from a statistical textbook or a **computer-generated sequence** (Table 4.3) (Beller et al. 2002).

4.1.4.2 Restricted Randomization

Restricted randomization procedures control the probability of obtaining an allocation sequence with an undesirable sample size imbalance in the intervention groups. In other words, if researchers want treatment groups of equal sizes, they should use restricted randomization (Schulz and Grimes 2002c).

A common type of restricted randomization is blocking. **Random permuted blocks** assign participants to treatments by chance but guarantee that equal numbers are assigned to each group as the trial progresses. Instead of randomly allocating all study subjects (simple randomization), this process randomizes participants in a series of blocks of specified sizes, such as 6 or 8 participants. For example, with a block size of 6 participants, 20 different permutations of three assignments to A and three to B exist (e.g., AAABBB, ABBBAA, AABBAB, etc.). Each of these blocks is given a number or range of numbers, and the sequence of 6-participant blocks is chosen with a random number table or computer-generated sequence of numbers. Thus, after every sixth participant (e.g., at 6, 12, 18, 24, etc.), equal numbers will have been assigned to A and to B. The longer the block, the less likely inquisitive investigators or study staff are to be able to decipher the block length and thus the upcoming assignment. In addition, randomly varying the block length and inserting a run of

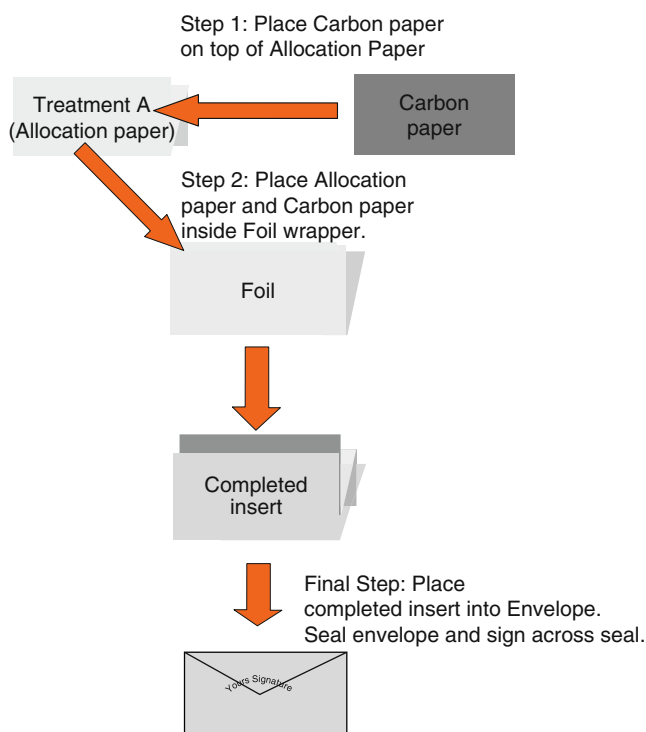


Fig. 4.1 Preparation of n envelope insert. *Step 1: initial preparation:* Cut the aluminum foil into n sheets that are of the same width as and twice the height of the envelope. The carbon paper should be cut into n envelope-sized sheets. Separate the n sheets of standard-size paper into 2 sets of $n/2$ sheets. On one set of $n/2$, print or write *Treatment A*, and on the second set, print or write *Treatment B*. *Step 2a: preparing treatment A envelopes:* Select 1 sheet of standard-sized paper marked *Treatment A* and fold to fit the envelope. Next, place 1 sheet of carbon paper on top of the folded treatment A allocation paper with the carbon side facing the paper (**Fig. 4.1**, Step 1) and fold 1 sheet of foil over both sides of the carbon-treatment A paper combination (**Fig. 4.1**, Step 2). Place

the completed insert into a blank envelope, with the carbon paper closest to the front of the envelope. If the completed insert is placed into the envelope properly, the double foil wrapper ensures that the envelope is truly opaque and cannot be read by holding it up against a strong light source. Complete all $n/2$ treatment A envelopes, seal each envelope and sign your name, in pen, over the top of the envelope seal. *Step 2b: preparing treatment B envelopes:* Prepare the treatment B envelopes as in step 2a. Do not mix treatment A envelopes with treatment B envelopes and do not write on the envelopes, except for signing your name over the seal. (Doig and Simpson 2005. Reprinted with permission from Elsevier)

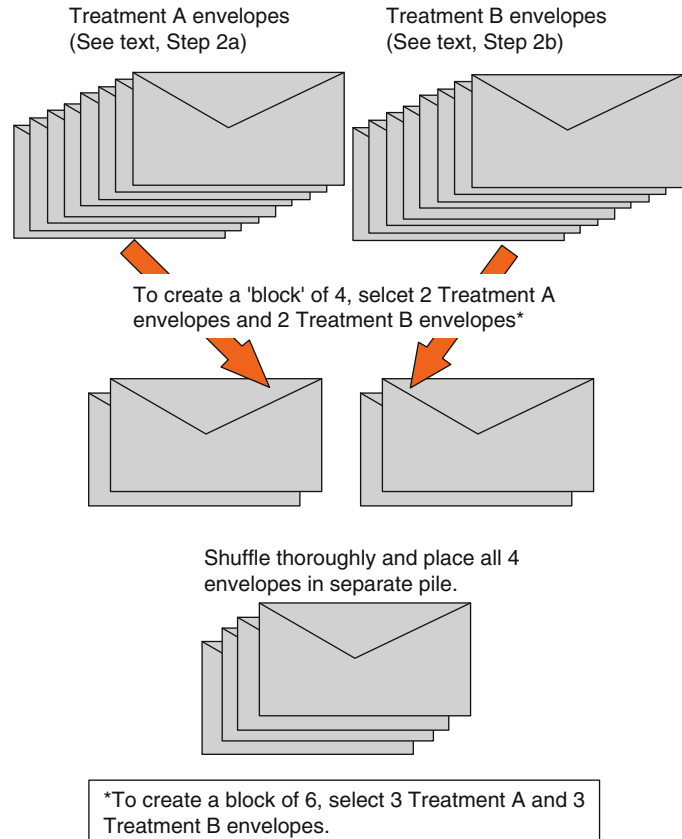
simple randomization (a process called mixed randomization) can render the allocation sequence nearly impenetrable (**Figs. 4.1 and 4.2**) (Grimes and Schulz 2002).

Biased-coin designs achieve much the same objective as blocking but without forcing strict equality (Wei and Lachin 1988). They therefore preserve most of the unpredictability associated with simple randomization. Biased-coin designs alter the allocation probability during the course of the trial to rectify imbalances that might be happening. Adaptive biased-coin designs, with the urn

design being the most widely studied, alter the probability of assignment based on the magnitude of the imbalance (Schulz and Grimes 2002c).

Replacement randomization repeats a simple randomization allocation scheme until a desired balance is achieved. Trial investigators should establish objective criteria for replacement. Although replacement randomization seems somewhat arbitrary, it is adequate as long as it is implemented before the trial begins. Moreover, it is easy to implement, ensures reasonable balance, and yields unpredictability.

Fig. 4.2 Permuted block randomization (blocks of 4 and 6) (Doig and Simpson 2005. Reprinted with permission from Elsevier)



The main limitation of replacement randomization is that it cannot ensure balance throughout the trial for interim analyses. Though rarely used, this approach emerged as the earliest form of restricted randomization (Schulz and Grimes 2002c).

4.1.4.3 Mixed Randomization

For mixed randomization, solution mixes simple randomization with permuted-block randomization. Simple randomization contributes the unpredictability to the approach, whereas permuted-block randomization contributes the balance. The mixed approach begins with an uneven block generated by a replacement randomization procedure (Table 4.4). Then, in its simplest form, standard permuted blocks of varying size follow. The replacement randomization sequence would establish inequality initially and make any anticipation of assignments improbable throughout the remainder of the trial (Schulz and Grimes 2002e).

4.1.4.4 Stratified Randomization

Randomization can create chance imbalances on baseline characteristics of treatment groups. Investigators sometimes avert imbalances by use of **pre-randomization stratification** on important prognostic factors, such as age or disease severity. In such instances, researchers should specify the method of restriction (usually blocking). To reap the benefits of stratification, investigators must use a form of restricted randomization to generate separate randomization schedules for stratified subsets of participants defined by the potentially important prognostic factors. Stratification without restriction accomplishes nothing—i.e., placebo stratification. Stratification in trials is methodologically valid and useful, but theoretical and pragmatic issues limit its use to those planning new trials (Grimes and Schulz 2002).

Stratification might be useful in small trials in which it can avert severe imbalances on prognostic factors. It will confer adequate balance (on the

Table 4.4 Mixed randomisation steps (Schulz and Grimes 2002e) (Reprinted from The Lancet 2002;359(9310):966–70 with permission from Elsevier)

Step 1: Generate one uneven block by replacement randomisation for the first participants
- Identify block size for the first uneven block. The block size can be odd or even and of any reasonable size, but usually in the range of 5–16. - Select a prespecified inequality in the sample sizes of the allocated groups for the first uneven block. - Generate a simple randomisation sequence (e.g., with a table of random numbers or a computer random number generator). - Inspect resultant sequence of assignments for matching or exceeding the desired prespecified inequality from Step B above.
If sufficiently unequal distribution of As and Bs, proceed to step 2; if not, go back to Step C above (iterate).
Step 2: Generate random permuted blocks for subsequent participants
- Select block sizes for the permuted blocks. Longer block sizes, such as 10–20, are more unpredictable than shorter block sizes, such as two to four. Longer block sizes should be preferred, unless an investigator needs approximate balance in a small trial or a small stratum of a trial. For example, an investigator might select block sizes of 8, 10, 12, and 14 as options. - Generate random permuted blocks, randomly varying the block size, as described in many texts. - Decide if an additional uneven block or simple randomisation block is to be interspersed in the trial. If not, complete the required sample size with random permuted blocks. Otherwise, identify a point at which to interject an uneven block or a simple randomised sequence. - If interjecting another uneven block by replacement randomisation, proceed back to step 1. If interjecting a simple random sequence, proceed to step 3.
Step 3: Generate a simple random sequence for interjection after a set of permuted blocks
- Identify the size of this simple random sequence. The size can be odd or even and of any reasonable size, but usually in the range of. We suggest an odd number to ensure some imbalance. - Generate a simple random sequence of the chosen size as suggested earlier. - Proceed to step 2, B.

stratified factors) and probably slightly more statistical power and precision. The added complexity of stratification yields little additional gain in large trials since randomization creates balanced groups anyway. Moreover, if imbalance arises, then investigators can statistically adjust on those prognostic variables (preferably pre-planned). Thus, stratification in large trials offers negligible advantages coupled with important, pragmatic disadvantages. Note one important exception; however, stratification by centre in multicentre trials promises some benefit with no added complexity to the trial implementers within each centre. Also, another potential exception arises in large multicentre trials in which investigators use central randomization for implementation of the sequence (Grimes and Schulz 2002).

4.1.4.5 Inappropriate Randomization Methods

Methods of allocation such as alternate allocation to treatment group or methods based on patient characteristics such as date of birth, order of entry

into the clinic, or day of clinic attendance are not reliably random. Such allocation sequences are predictable and not easily concealed, thus reducing the guarantee that allocation has indeed been random and that no potential subjects have been excluded by foreknowledge of the intervention (Beller et al. 2002).

4.1.4.6 Allocation Concealment

Allocation concealment is the second critical bias-eliminating element of randomization. This term means that both participants and trial staff are unaware of the upcoming treatment assignment (Grimes and Schulz 2002).

Blinding should not be confused with allocation concealment. Concealment seeks to prevent selection bias and protects the assignment sequence before and until allocation, while blinding seeks to prevent ascertainment bias and protects the sequence after allocation (Schulz and Grimes 2002b).

As with methods of random-sequence generation, methods of allocation concealment vary widely

in their rigor (Schulz and Grimes 2002b). The following techniques are considered adequate:

- Centralized or remote randomization schemes
- Randomization schemes controlled by a pharmacy
- Numbered or coded containers in which capsules from identical-looking, numbered bottles are administered sequentially
- On-site computer systems, where allocations are in a locked unreadable file
- Sequentially numbered opaque, sealed envelopes (Higgins and Green 2011; Fenwick et al. 2008)

Failed Concealment from the Investigator or Clinician: Single-centre trials in which randomization is conducted on site and the randomization method is known to the participating investigators are at high risk of this problem. If an investigator has influence over the number of patients enrolling in a trial, or the order in which they are randomized, he or she has the potential to introduce selection bias into the study by directing particular patients into preferred treatment groups while excluding others. Ideally, a successful allocation concealment process is ensured when all investigators are ignorant of future treatment allocations and have no control over the order of patients randomized into the trial (Forder et al. 2005).

Failed Concealment from the Patient: Patients may change their willingness to participate in a trial if they know or suspect which treatment they will receive. If patients are aware of their allocation before randomization, they may be influenced to withdraw before randomization or wait until their preferred treatment is available before entering the study (Forder et al. 2005).

Comparisons of adequately and inadequately concealed allocation in randomized trials of the same intervention provided high-quality evidence that concealment can be crucial in achieving similar treatment groups and therefore, unbiased estimates of treatment effects (Kunz et al. 2007). Studies with inadequate concealment tended to overestimate treatment effects (Kunz et al. 2007; Odgaard-Jensen et al. 2011).

A meta-analysis of these observational studies showed that improper allocation concealment

produced the largest impact on size of treatment effect (Jüni et al. 2001). Treatment was found to be 30% more beneficial if concealment was inadequate or unclear (combined odds ratio 0.70; 95% CI, 0.62, 0.80). When the robustness of conclusions drawn from meta-analyses to exclusion of randomized trials without reported adequate allocation concealment was investigated, two-thirds of conclusions in favour of one of the interventions were no longer supported if only trials with adequate allocation concealment were included. The loss of support mainly reflected loss of power (the total number of patients was reduced by 49%) but also a shift in the point estimate towards a less beneficial effect (Pidal et al. 2007).

Investigations of the dental literature (Antczak et al. 1986; Montenegro et al. 2002; Sjögren and Halling 2002) have consistently indicated that reported methods and use of allocation concealment and examiner masking are not optimal and are similar comparing these fields (Fenwick et al. 2008). Montenegro et al. (2002) showed that although 91% of trials in periodontology were described as randomized, adequate methods for randomization and allocation concealment were found in 17% and 7% of studies, respectively. However, insufficient evidence to support or refute an effect of allocation concealment or examiner masking on the magnitude of the treatment effect in periodontology was found. Retrospective power calculations suggest that at least 265 trials would be needed to demonstrate an effect if the magnitude of the overestimation of the magnitude of clinical effect was 0.5 mm. Future definitive research on this topic will therefore need to examine a much larger sample of trials (Fenwick et al. 2008). When the quality of randomized controlled trials (RCTs) concerned with the effectiveness of oral implants was assessed, it was showed that randomization and concealment allocation procedures were not described in 30 out of 43 articles (70%) (Esposito et al. 2001).

Reports of randomized controlled trials should document that the randomization yielded groups similar in all important respects, except for the exposure being studied. The first table of such reports usually fulfils this role. This table presents demographic and other relevant clinical fea-

tures by treatment group, for example, age, race, and parity. This enables readers to know the types of participants enrolled. Can the results be extrapolated to the readers' own patients? In addition, are the groups similar in all important respects except for the exposure being studied? (Grimes and Schulz 2002).

4.1.4.7 Minimization

Minimization is an assignment strategy, similar in intention to stratification, that ensures excellent balance between intervention groups for specified prognostic factors. The next participant is assigned to whichever group would minimize the imbalance between groups on specified prognostic factors. Minimization is an acceptable alternative to random assignment (CONSORT Glossary 2007).

Descriptions of Randomization. Examples

- *Assignment to test (LDD) or control (inactive identical placebo) group was random. The sequence was computer generated in blocks of four by the research coordinator. Randomization was concealed by the use of sequentially numbered, identical containers containing active or identical-appearing placebo medication. The randomization code was held centrally by the research coordinator remote from the study and was not broken until completion of the data analyses (Needleman et al. 2007).*
- *Patients were assigned consecutive and ascending numbers at the enrolment visit. Prior to the beginning of active therapy, each subject was randomized to a single treatment group (test or control) by using the method of randomly permuted blocks (Varela et al. 2011).*
- *A randomized, evaluator-blinded, controlled trial with two different active treatments was performed. Participants were randomly allocated to an individually tailored oral health educational programme (experimental group) or to a standard treatment programme (control group). The randomization was made in blocks of various sizes by a random computer table. Allocation concealment was secured by (1) having a person not involved with the clinic perform the randomization i.e. neither*

the examiner nor the therapist could influence the allocation of group belongings and (2) providing the dental hygienists with sealed consecutively numbered envelopes containing only the assignment for an individual subject. The dental hygienist had not met the patient before the assignment (Jönsson et al. 2010).

4.1.5 Blinding of Treatment

Blinding is used in combination with randomization and has several **benefits**: usually reduces differential assessment of outcomes (information bias) but can also improve compliance and retention of trial participants while reducing biased supplemental care or treatment (sometimes called co-intervention) (Table 4.5) (Schulz and Grimes 2002b).

Some people prefer the term **masking** to blinding to describe the same procedure. Blinding, however, conveys a strong bias prevention message. The imagery of blindfolding, a total covering of the eyes, conveys stronger bias prevention than masking, where eye holes could permit viewing. Blinding also suggests a more secure procedure to some and permeates the ICH guidelines (The International Conference on Harmonization (ICH 2000) is an intensive tripartite collaboration between regulatory authorities in Europe, Japan, and the USA to develop common guidelines for the design, implementation, and reporting of clinical trials) (Schulz and Grimes 2002b).

Blinding is not well understood. It was demonstrated an important variation among physicians in their interpretation of who is unaware of allocation when authors use the terms 'single', 'double', and 'triple' blind. This variability suggests that the **current terminology** hinders readers who are trying to accurately assess trial validity. The variation in textbook definitions demonstrates that there is no consensus among authorities as to what these terms mean. A survey of 200 recent RCTs in high-impact medical journals demonstrated that, in 50% of trials, authors who use the term 'double blind' failed to make any statement about who was blinded and specified only one group—likely omitting information about one or more other

Table 4.5 Potential benefits accruing dependent on those individuals successfully blinded (Schulz and Grimes 2002b)
Reprinted from *The Lancet* 2002;359(9307):696–700 with permission from Elsevier)

Individuals blinded	Potential benefits
Participants	Less likely to have biased psychological or physical responses to intervention More likely to comply with trial regimens Less likely to seek additional adjunct interventions Less likely to leave trial without providing outcome data, leading to lost to follow-up
Trial investigators	Less likely to transfer their inclinations or attitudes to participants Less likely to differentially administer co-interventions Less likely to differentially adjust dose Less likely to differentially withdraw participants Less likely to differentially encourage or discourage participants to continue trial
Assessors	Less likely to have biases affect their outcome assessments, especially with subjective outcomes of interest

groups who were blinded—in another 35%. Physician respondents identified 10, 17, and 15 unique interpretations of single, double, and triple blinding, respectively, and textbooks provided 5, 9, and 7 different definitions of each. The frequencies of the most common physician interpretation and textbook definition were 75% (95% confidence interval [CI], 65–83%) and 74% (95% CI, 52–90%) for single blinding, 38% (95% CI, 28–49%) and 43% (95% CI, 24–63%) for double blinding, and 18% (95% CI, 10–28%) and 14% (95% CI, 0–58%) for triple blinding, respectively (Devereaux et al. 2001).

Recently, the reporting of blinding in protocols, and its relation to the reporting of blinding in publications, has been studied (Hróbjartsson et al. 2009). Protocols focusing on the methodological aspects of trials are formulated before results accumulate and have no space restrictions. The information on blinding in protocols could, therefore, be more comprehensive and reliable than the information provided in publications and from contact to authors. Publications generally conveyed less information on blinding than protocols. Overt contradictions between the reporting of blinding in protocols and publications were rare. In most trials, even the combined information from protocol and publication was inadequate for a clear identification of the blinding status of key trial persons and the mechanisms involved in initiating and maintaining blinding (Table 4.6). In the publications for 52 of the 58 trials for trials labelled as ‘double blind’, it was unclear whether

participants, health-care providers, and data collectors had all been blinded. In 2 of these 52 trials (4%), the protocol clarified that all three key trial persons had been blinded. The protocols clarified that blinding had been planned for participants in 66% of trials, for health-care providers in 39% of trials, and for data collectors in 8% of trials. Lack of reporting of the blinding of a key trial person in a publication reflected undisclosed lack of blinding as defined in the protocol in 16% of the trials. The reporting on blinding in protocols is more comprehensive than in publications, but a large proportion of protocols still report blinding unclearly (Hróbjartsson et al. 2009).

Kolahi and Abrishami (2009) suggested a new blinding protocol acting as a gold standard for design and conduction of unbiased clinical trials. All of the following five **categories of key trial persons should be blinded to treatment allocation**: (a) participants, (b) health-care providers (e.g., doctor or nurses responsible for care), (c) data collectors (e.g., the auxiliary staff or doctor responsible for measurement or collection of outcome data), (d) outcome assessors (e.g., the data monitoring board, endpoint committee, or judicial assessors of outcome), and (e) data analysts. Furthermore, personnel writing the manuscript should be blinded to treatment allocation. The method implies that two manuscripts are written; the first assumes treatment A to be the experimental treatment and treatment B to be the control; the second manuscript is based on the reverse assumption, that is, that treatment A is the

Table 4.6 Concordance in the reporting of blinding in corresponding protocols and publications of randomized trials (N=73) (Hróbjartsson et al. 2009) (Reprinted with permission from Elsevier)

Concordance	n (%)
<i>Blinding reported with global term</i>	
Match (same term used in protocol and publication)	55 (75)
‘Double blind’ (n=49)	
‘Blind’ (n=1)	
‘Open’ ^a (n=1)	
No global term used ^a (n=4)	
Contradictions	2 (3)
‘Double blind’ in protocol; ‘single blind’ in publication (n=1)	
‘Open’ in protocol; ‘double blind’ in publication (n=1)	
Discrepancies	16 (22)
‘Double blind’ in protocol; no term or ‘blind’ in publication (n=5)	
‘Single blind’ in protocol; no term in publication (n=1)	
‘Blind’ in protocol; no term or ‘single blind’ in publication (n=2)	
‘Open’ in protocol; no term in publication ^a (n=2)	
No term in protocol; ‘double blind,’ ‘single blind,’ ‘blind,’ or ‘open’ in publication (n=6)	
<i>Blinding of individual key trial persons reported</i>	
Match (same key trial persons reported as blinded in protocol and publication)	23 (32)
Participants (patients or volunteers) (n=2)	
Data collectors (n=2)	
Data collectors and judicial outcome assessors (n=1)	
No key trial person described (n=18)	
Contradictions	1 (1)
Data analysts reported as not blinded in protocol, but as blinded in publication (n=1)	
Discrepancies	49 (67)
Additional key trial persons reported in protocol (n=38)	
Additional key trial persons reported in publication (n=16)	
Additional key trial persons reported in both protocol and publication (n=5) ^b	

^aBlinding of a key trial person was clearly described

^bFor example, patients were reported as blinded in the protocol (but not in the publication), and health-care providers were reported as blinded in the publication (but not in the protocol)

control treatment. Both manuscripts must be completed and approved by the authors before the code is broken (Kolahi and Abrishami 2009).

Placebos are often used to maintain blinding. A placebo is a pharmacologically inactive drug or agent that is given to the control group in a trial. Administration of a placebo to the control group balances the psychological placebo effect that occurs in the experimental group. In addition, it maintains blinding. Use of a placebo is appropriate when no established therapy exists for a condition; if an effective treatment is available, then use of a placebo is inappropriate and unethical. Instead, the new treatment can be compared with an existing treatment. When used, pla-

cebos should be identical in appearance, smell, and taste to prevent deciphering of the treatment assignments. When disparate interventions are being compared (e.g., an injection vs. a tablet), placebos can still be used to maintain blinding. In this situation, a **double-dummy** approach is helpful: each participant receives one active and one inactive treatment. For example, those assigned to the injection group would receive an active injection and a placebo tablet, whereas those allocated to the tablet group would get a placebo injection and an active pill (identical to the placebo tablet) (Schulz and Grimes 2002b).

A classification scheme of **blinding techniques** in pharmacological and non-pharmaco-

logical treatments reviewed (Boutron et al. 2006, 2007) was developed to distinguish between (1) the blinding of key trial person—participants, health-care providers, or other caregivers—that rely on the category of treatment (surgical/technical procedure, participative interventions, devices) and comparator assessed (Fig. 4.3), and (2) the blinding of outcome assessors depending on the primary outcome considered (Fig. 4.4). Although other classification systems are possible, the presented one focuses on the category of treatment, the comparator, and the primary outcomes, believing them to be more practical for those planning RCTs that involve blinding.

Several **methods used to assess the success of blinding** were reported in the literature (Kolahi et al. 2009):

- The ‘2×3 format’: Subjects could be asked to guess their treatment assignment, but they may be allowed to express treatment guess or uncertainty, that is, subjects’ guessing options could be ‘active versus placebo (or control) versus do not know’ (Lao et al. 1999).
- The ‘2×5 format’: The certainty of the guess about the subjects’ treatment can be rated on a 5-point scale: (1) strongly believe the treatment is active, (2) somewhat believe the treatment is active, (3) somewhat believe the treatment is placebo, (4) strongly believe the treatment is placebo, and (5) do not know (Boutron et al. 2005; LaRosa et al. 1994).
- Pretrial evaluation of the potential for unblinding by conducting a test on an independent panel of volunteers who are not participating in the actual trial, in a preliminary study by asking them to try to distinguish different matching preparations for active versus control treatments (Walter et al. 2005).
- Inter-group comparison of the change in beliefs (and not necessarily their correctness) between the start and end of a trial. This method shifts the focus away from the correctness of beliefs regarding treatment assignment; instead, it emphasizes the beliefs themselves and their change pattern during the trial (Rees et al. 2005).

The optimal **timing and frequency of blinding assessments** is a controversial issue. Kolahi

et al. (2009) recommended assessments of blinding before the trial by the third party for independent credibility, in the early stage of the trial with great caution for credibility at specific trial setting, and at the end of the trial for overall assessments of trial participants.

Clearly, a vast majority of blinded trials do not report test or sufficient **evidence for the success of blinding**. In medical literature, 31 out of 1,599 trials (2%) reported tests for the success of blinding. Test methods varied, and reporting was generally incomplete (Hrobjartsson et al. 2007). Similar results were reported by Fergusson et al. (2004) who revealed that only 7 of the 97 (7%) general medicine trials provided evidence on the success of blinding, with five reporting that the success of blinding was imperfect. In trials from psychiatric journals, the success of blinding was reported in 8 of the 94 trials, with four reporting that the blinding was imperfect. Overall, only 4 of the 191 (2%) trials assessed blinding in the participants and either the outcome assessors or the investigators (Fergusson et al. 2004).

To deal with data from surveys on treatment guessing, James et al. (1990) and Bang et al. (2004) developed **blinding indices (BIs)**. James et al.’s method is a variant of the kappa statistic, which focuses on disagreement between treatment allocation and guess. Bang et al.’s BI estimates the percentage of correct guessing beyond the level expected by chance in each treatment group. A unique feature in blinding assessment is how to deal with ‘do not know’ answer. Since their introduction, both BIs have been used in several studies to quantitatively measure the level of correct guess, a surrogate indicator of unblinding (Park et al. 2008). The blinding index proposed by Bang et al. (2004) is scaled to an interval of −1 to 1, 1 being complete lack of blinding, 0 being consistent with perfect blinding, and −1 indicating opposite guessing which may be related to unblinding. It has the ability to detect a relatively low degree of blinding, response bias, and different behaviours in two arms. These two BIs are based on different paradigms and assumptions, and naturally, they both have their own strengths and weaknesses. Computing modules for the two BIs are available for Stata (StataCorp

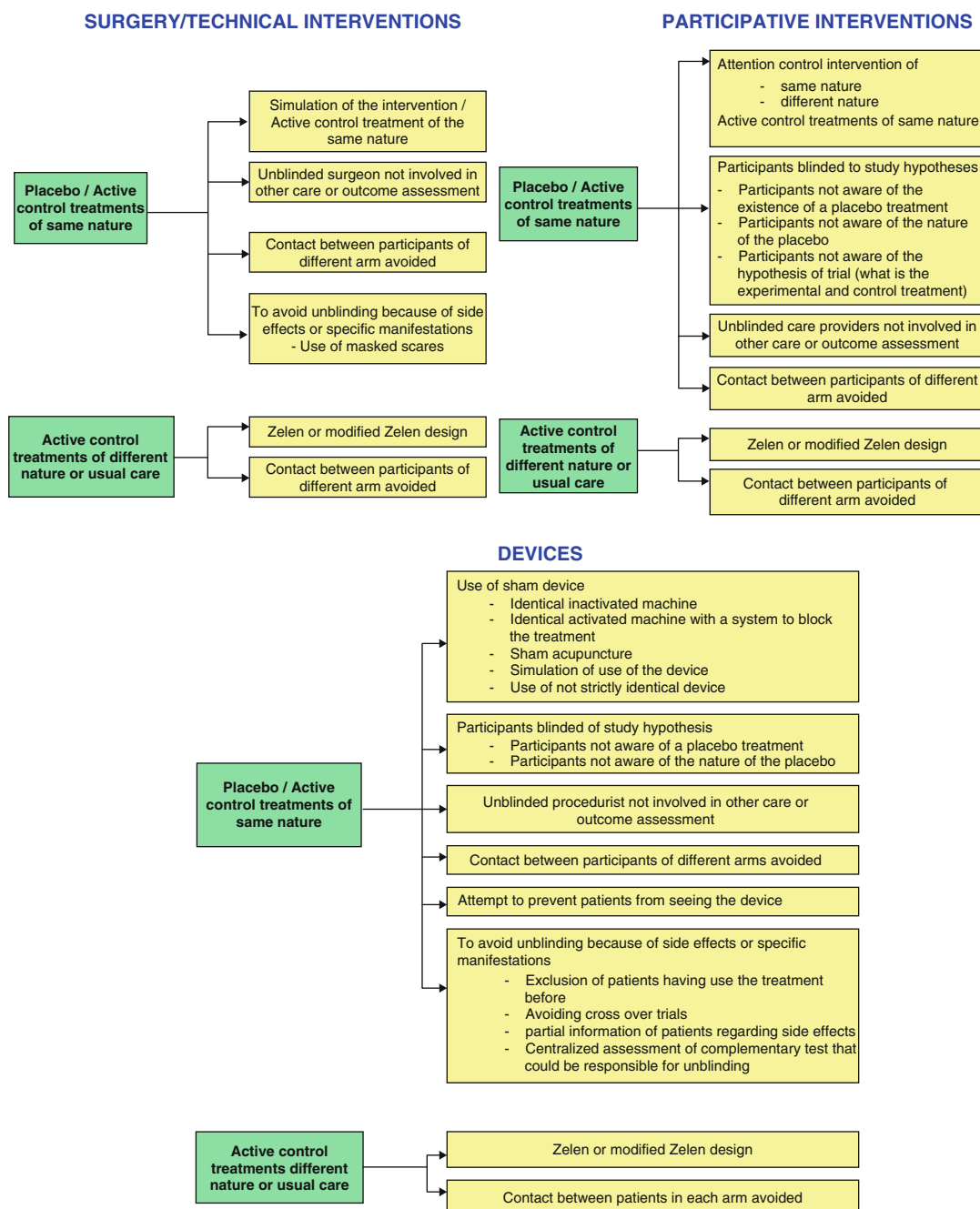


Fig. 4.3 Methods of blinding participants, health-care providers, or other caregivers that rely on the category of treatment and comparator assessed (Boutron et al. 2007. doi:10.1371/journal.pmed.0040061)

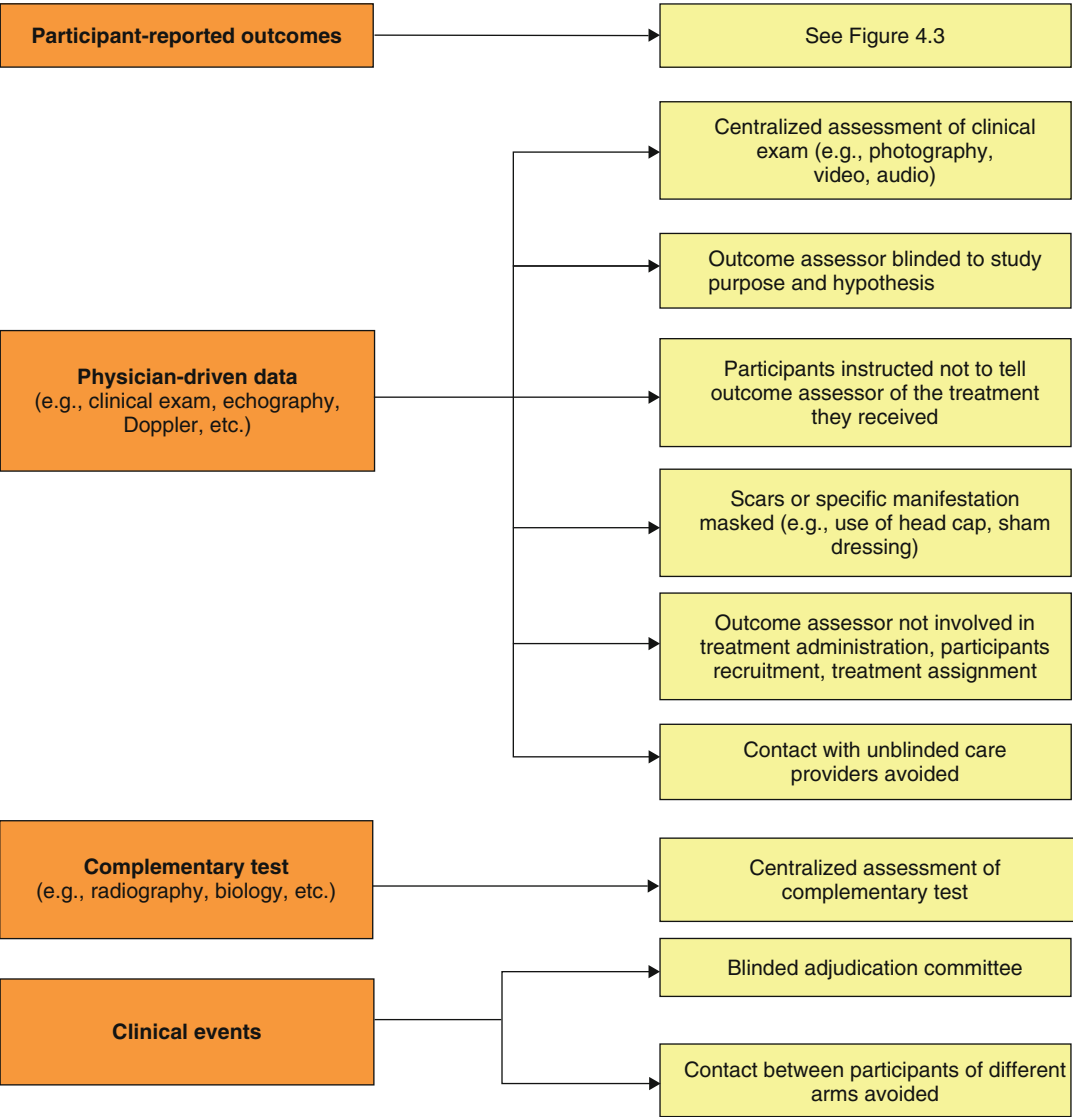


Fig. 4.4 Methods of blinding outcome assessors depending on the primary outcome considered (Boutron et al. 2007. doi:10.1371/journal.pmed.0040061)

LP, College Station, TX, USA) since March 2008, which are available at <http://ideas.repec.org/c/boc/bocode/s456898.html> (Kolahi et al. 2009). Several studies have examined the effects of masking on quality assessments of clinical trials (Jadad et al. 1996; Moher et al. 1995; McNutt et al. 1990). The results show little consistency in direction or magnitude. The impact of blinding within a meta-analysis was variable (Moher et al.

1998). Masked assessments provided significantly higher quality of trial scores than unmasked assessments (mean 2.74 [SD 1.10] vs. 2.55 [1.20]). Low-quality trials, compared with high-quality trials, were associated with an increased estimate of benefit of 34% (ratio of odds ratios [ROR] 0.66 [95% CI 0.52–0.83]). Colditz et al. (1989) analysed 113 reports published in 1980 in a sample of medical journals to relate features of study design to the magnitude of gains attributed

Table 4.7 Assessment of blinding success in the study of LaRosa et al. (1994): main data (upper) and ancillary data (lower) for ‘do not know’ responses (Kolahi et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Intervention	Subject’s guess, <i>n</i> (%)						Question could not be asked, <i>n</i> (%)
	1	2	3	4	5	Total	
Active	38 (9)	44 (10)	170 (40)	21 (5)	4 (1)	277 (66)	0
Placebo	11 (2)	16 (4)	83 (20)	21 (5)	8 (2)	139 (33)	0
Total	49 (12)	60 (14)	253 (61)	42 (10)	12 (3)	416 (100)	0

Intervention	Subject’s guess, <i>n</i> (%)			Question could not be asked, <i>n</i> (%)
	Active	Placebo	Total	
Active	79 (31.2)	86 (33.9)	170 (67.1)	5 (1.9)
Placebo	36 (14.2)	45 (17.7)	83 (32.8)	2 (0.7)
Total	115 (45.4)	131 (51.7)	253 (100)	7 (2.7)

1 strongly believe the treatment is active, 2 somewhat believe the treatment is active, 3 do not know, 4 somewhat believe the treatment is placebo, 5 strongly believe the treatment is placebo

to new therapies over old. Randomized controlled trials that did not use a double-blind design had a higher likelihood of showing a gain for the innovation than did double-blind trials. In contrast, in a recent meta-analysis, the multilevel meta-regression analysis using the entire data set demonstrated that neither concealment of allocation ($P=0.96$) nor blinding ($P=0.32$) was a significant predictor of the differences in treatment effect size (Bassler et al. 2010).

Reporting of the Results. As suggested in the CONSORT guidelines (Moher et al. 2001; Altman et al. 2001), investigators should not use only the single-blind, double-blind, or triple-blind terminology but should also explicitly state who was blinded and how. Moreover, if the researchers contend that the trial investigators, participants, and assessors were blinded—that is, double blind—then they should provide information about the mechanism (capsules, tablets, film, &c), similarity of treatment characteristics (appearance, taste, administration), and allocation schedule control—for example, location of the schedule during the trial, when the code was broken for the analysis, and circumstances under which the code could be broken for individual instances. Such additional information can lend support to or undermine claims of double blinding (Schulz and Grimes 2002b). Kolahi et al. (2009) suggested that subjects’ answers could be reported for relevant parties of the five categories

of key trial persons who were tested for the success of blinding. Also, subjects’ answers could be reported for each trial arms. It was recommended for investigators to report blinding data in 2×3 or 2×5 tables or adequate details that allow readers to reconstruct one of these tables. **Table 4.7** in the study of LaRosa et al. (1994) is a good example for the 2×5 table.

When the explicit reporting of blinding status of each of the groups involved in RCTs published in five high-impact medical journals was evaluated in 2002, of the 160 RCTs in journals adopting the CONSORT Statement and 40 RCTs published in the non-adopter journal, 16% and 10% explicitly reported blinding of participants, 4% and 8% of health-care workers, 11% and 13% of data collectors, 21% and 30% of outcome assessors, 3% and 0% of data analysts, and 0% and 0% of manuscript writers, respectively. Although only one of the five studies described as single blind did not identify which ‘single’ group was blind, two of these studies identified more than one group as blind to allocation. In these ‘single-blind’ studies, the groups clearly identified as blind to allocation did not include participants or health-care providers. Rather, these studies described data collectors, outcome assessors, and data analysts as blind. Eighty-three reports used the term ‘double blind’. These RCTs provided eight unique combinations of two or

more groups that remained blind in double-blind trials (Montori et al. 2002).

Among 200 blinded randomized clinical trials published in 2001 randomly sampled from the Cochrane Central Register of Controlled Trial, 78% articles described trials as ‘double blind’ but in only 2% of such articles the blinding status of patients, health-care providers, and data collectors was explicitly described, while 56% articles did not describe the blinding status of any trial person, and 26% articles reported no blinding relevant information at all beyond the trial being ‘double blind’. Survey responders provided 15 different operational meanings of the term ‘double blind’ and typically felt that their preferred definition was the most widely used: patients were blinded in 97% ‘double-blind’ trials and health-care providers in 89%, while 19% ‘double blind’ trials had not blinded either patients, health-care providers, or data collectors (Haahr and Hróbjartsson 2006).

The assessment of the quality of RCTs in periodontology, using evidence-based components, revealed that patient and caregiver blinding was absent in 43% and 57% of studies, respectively. Approximately one-quarter of studies provided inadequate information to determine whether these types of blinding had been used. Only 17% of studies positively demonstrated caregiver blinding. Examiner blinding was not possible for 16% of studies, mainly due to very clear differences in treatments between test and control groups. Out of the remaining studies, 55% had definite examiner blinding, 7% had no blinding, and 38% provided inadequate information to determine (Montenegro et al. 2002). No attempt at blinding was reported in 31 out of 43 RCTs studies (72%) concerned with the effectiveness of oral implants was assessed (Esposito et al. 2001).

Descriptions of Blinding. Examples

- *This was a parallel arm, randomized, identical placebo-controlled trial with masking of examiner, care-giver, participant and statistician, with 6 months of follow-up..... The test drug was 20 mg doxycycline (Periostat, CollaGenex Pharmaceuticals, Newtown, PA,*

USA) twice daily for 3 months. The placebo was identical, with the exception of omission of doxycycline (Needleman et al. 2007).

- *Subjects were randomly assigned to one of the two treatment regimens. The randomized allocation was carried out using a computer software program (SPSS Inc., Chicago, IL, USA). When performing their respective study tasks, the study examiner and the therapist were not jointly present with the study subject. Study subjects were instructed not to discuss the therapy with the study examiner. The study examiner was, therefore, unaware of study treatment allocation. The study examiner performed all clinical measurements and collected samples for microbiological analysis. The clinician performing treatment had 47 years of clinical experience in the mechanical treatment of implants with a diagnosis of periimplantitis. The treatment code was not revealed until all microbiological assays had been completed and the electronic file (SPSS data file) had been established (Persson et al. 2010).*
- *An external agent randomized the treatments, by two computer-generated lists, one for each centre, by coding identical bottles with either test or placebo mouth rinses with consecutive numbers. Numbers were assigned to patients consecutively. Patients were stratified in two categories according to their tobacco habit as: non-smokers (including non-smokers and smokers of <10 cigarettes/day) and smokers of 10 or more cigarettes/day. Codes were not revealed until the study was finished. Both the examiners and the subjects were blinded to the content of the bottles. No attempt to blind examiners for tooth staining was made.... The placebo rinse was identical, except that it lacked the active agents, chlorhexidine and cetyl-pyridinium chloride (Escribano et al. 2010).*
- *Medication and placebo were prepared and encased in identical opaque coded bottles by the School of Pharmacy at the Health Sciences Center/UFRJ, according to the computer-generated list. Allocation was implemented by a senior investigator (M.C.M.B.T.), not directly involved with the examination or*

Table 4.8 Baseline characteristics of subjects (Needleman et al. 2007) (Reprinted with permission from John Wiley & Sons, Inc.)

	<i>Subject means (95% CI)</i>		
	<i>Placebo (N=18)</i>	<i>LDD (N=16)</i>	<i>Difference</i>
<i>Subjects</i>			
<i>Smoking (years)</i>	26.1 (13.0, 39.1)	24.4 (8.0, 40.9)	1.6 (−3.6, 6.8)
<i>Cigarettes (per day)</i>	18.3 (7.0, 29.5)	16.7 (3.2, 30.2)	1.6 (−2.7, 5.9)
<i>CO^a (% volume)</i>	5.0 (−0.3, 10.2)	4.6 (0.0, 9.2)	0.4 (−1.3, 2.1)
<i>CO^a (ppm)</i>	27.7 (−0.8, 56.1)	25.3 (−2.4, 52.9)	2.4 (−7.2, 12.1)
<i>Cotinine (ng/ml)</i>	379.6 (0.0, 771.6)	327.2 (19.5, 634.9)	52.4 (−69.9, 174.7)
<i>Treated teeth/sites</i>			
<i>PPD (mm)</i>	5.79 (4.93, 6.65)	5.96 (4.92, 7.01)	−0.17 (−0.51, 0.16)
<i>REC (mm)</i>	−0.12 (−2.13, 1.89)	0.28 (−1.82, 2.39)	−0.41 (−1.12, 0.30)
<i>CAL (mm)</i>	5.67 (3.52, 7.82)	6.25 (3.99, 8.50)	−0.58 (−1.34, 0.18)
<i>BoP^a (%)</i>	76.3 (52.5, 100.1)	65.8 (21.7, 109.8)	10.5 (−2.1, 23.1)
<i>Plaque^a (%)</i>	76.8 (42.9, 110.8)	79.5 (37.6, 121.4)	−2.7 (−16.0, 10.6)
<i>ICTP^b (pg/site)</i>	2.25 (0.00, 4750.76)	0.81 (0.00, 9279.67)	1.44 (−1.95, 3.99)

CI confidence interval, ICTP carboxyterminal telopeptide of type I collagen, BoP bleeding on probing, CAL clinical attachment level, CO carbon monoxide, LDD low-dose doxycycline

^aNot normally distributed hence means and 95% CIs have limited meaning

^bGeometric mean (i.e., natural logarithm taken before mean is calculated and inverse logarithm applied afterwards)

treatment procedure. Identification code was kept concealed from all individuals directly involved in the study, until the statistical analysis was carried out (Varela et al. 2011).

4.1.6 Baseline Comparisons

Although randomization eliminates systematic bias, it does not necessarily produce perfectly balanced groups with respect to prognostic factors. Differences due to chance remain in the intervention groups—that is, chance maldistribution. Statistical tests, however, account for these chance differences. The process of randomization underlies significance testing and is independent of prognostic factors, known and unknown. Nevertheless, researchers should present distributions of baseline characteristics by treatment group in a table. Such information describes the hypothetical population from which their trial arose and allows readers to see the possibilities of generalization to other populations. Furthermore, it allows physicians to infer the results to particular patients (Schulz and Grimes 2002a).

Examples

- *The age range of the participants was 32–50 years. The placebo group was slightly younger, with a mean birth year of 1957.8 [standard deviation (SD)56.1], than the Periostat group, with a mean birth year of 1960.3 (SD57.6), although both sexes were similarly aged with males' mean birth year being 1959.8 (SD56.6) and females' mean birth year being 1958.1 (SD57.3). Baseline characteristics are summarized in Table 4.8. Subjects within groups were generally well matched at baseline, with measures of tobacco exposure being similar except for salivary COT, which was slightly higher in the placebo group. Clinically, a 10% difference for BoP was also observed (Needleman et al. 2007).*
- *The baseline characteristics of the 113 participants receiving the individually tailored oral health educational programme or standard treatment are displayed in Table 4.9. Randomization was successful, as there was no statistically significant difference in the demographic variables or background characteristics between the groups (p-value not shown in tables) (Jönsson et al. 2009).*

Table 4.9 Baseline characteristics of the participants in the experimental and control group (Jönsson et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

	<i>Experimental (n = 57)</i>	<i>Control (n = 56)</i>
<i>Mean age</i>	52.4 (8.4)	50.1 (10.3)
<i>Cigarette smokers</i>	24 (42.1%)	20 (35.7%)
<i>Gender</i>		
<i>Female</i>	32 (56.1%)	28 (50.0%)
<i>Male</i>	25 (43.9%)	28 (50.0%)
<i>Marital status</i>		
<i>Married or cohabitants</i>	45 (78.9%)	40 (71.4%)
<i>Single</i>	12 (21.1%)	16 (28.6%)
<i>Education</i>		
<i>Elementary school</i>	14 (24.6%)	13 (23.2%)
<i>High school</i>	21 (36.8%)	23 (41.1%)
<i>University</i>	22 (38.6%)	19 (33.9%)
<i>Ethnicity</i>		
<i>Swedish</i>	46 (80.7%)	50 (89.3%)
<i>Other</i>	11 (19.3%)	6 (10.7%)
<i>Previous visits to dentist</i>		
<i>Once a year</i>	37 (64.9%)	37 (66.1%)
<i>Every second year</i>	10 (17.5%)	10 (17.9%)
<i>At irregular intervals</i>	10 (17.6%)	9 (16.0%)
<i>Previous visits to dental hygienist</i>		
<i>Several times per year</i>	20 (35.1%)	20 (35.7%)
<i>Once a year</i>	11 (19.3%)	21 (37.5%)
<i>At irregular intervals</i>	25 (43.9%)	15 (26.8%)
<i>Number of teeth</i>		
<i>Baseline</i>	25.3 (3.9)	25.0 (4.6)
<i>At the start of non-surgical treatment</i>	23.3 (4.0)	23.2 (4.6)

4.1.7 Sample Size and Power Calculations in RCTs

One of the most important aspects of study design is the estimation of required sample size. Proper consideration of the number of patients that need to be recruited will help ensure that a study has maximal efficiency to answer the primary question of interest. Studies that are designed with too few patients will almost certainly be unable to show a difference between treatments. Failure to appropriately anticipate the study size during the design phase can raise important ethical issues. An underpowered study having too few patients is unlikely to answer the primary study question, and information obtained from well-intentioned participants will not be useful. At the other extreme, an overpowered study enrolling more patients than necessary may needlessly subject some partici-

pants to an inferior treatment for longer than necessary (Scales and Rubenfeld 2005; Eng 2003).

Sample size considerations are not only important for clinical researchers. Clinicians who are trying to interpret and apply research findings require a basic understanding of the principles underlying sample size calculation. This is acknowledged in the revised Consolidated Standards of Reporting Trials guidelines, which recommend that a description be provided of the methods used to estimate sample size and when applicable, explanation of any interim analyses and stopping rules (Scales and Rubenfeld 2005). To help clinicians critically appraise design issues relating to study size and power in clinical research, Scales and Rubenfeld (2005) have proposed several questions that can be asked of every study (Table 4.10) and have presented several strategies that can decrease the required sample size for a study when the initial

Table 4.10 Questions for critically appraising sample size calculation for a clinical trial (Scales and Rubenfeld 2005) (Reprinted with permission from Elsevier)

1. Are the methods used to estimate the required sample size described?
2. What was the end point used to calculate the required sample size?
- Is the end point used to estimate the required sample size the same as the primary outcome reported in the study? - Is the choice of primary end point used to estimate sample size reasonable and clinically important? - Is a clear and acceptable definition of the primary end point provided? Is the time at which it will be measured specified?
3. For dichotomous end points, is the baseline (control) event rate specified and in keeping with existing knowledge and experience?
Is the predicted event rate similar to that which was actually observed during the trial?
4. For continuous variables, is the SE of measuring the primary outcome specified?
Is the predicted SE similar to that which was actually observed during the trial?
5. Is the minimal clinically important difference that the study is designed to detect specified?
- Is this minimal clinically important difference plausible given the anticipated treatment effect? - Would you, your colleagues, or the patient consider this minimal clinically important difference to be either too high or too low?
6. Are the rates of type I error specified and acceptably low?
If interim analyses have been conducted, have the cumulative effect on type I error probability been considered?
7. Are the rates of type II error specified and acceptably low?
Conventionally set at .2, or lower for efficacy studies, but should be much lower (i.e., b.1 or .05) if attempting to demonstrate equivalence or noninferiority.
8. Were 1- or 2-sided tests of significance used?
If tests were 1-sided, is there any possibility that the new treatment may have been harmful compared with the control intervention?
9. Did the researchers consider the potential for dropouts or missing data when estimating sample size?

estimate is impractically large (Table 4.11). It should be noted that most of these strategies are associated with trade-offs that can have important implications on study design and the generalizability of results. For example, improving the precision of measuring the primary endpoint can decrease SD and hence reduce the required sample size but can increase cost and time requirements. In contrast, increasing the tolerable type I error rate will certainly decrease the number of patients needed but should seldom be considered an acceptable strategy (Scales and Rubenfeld 2005).

4.1.7.1 Types I and II Errors

In clinical research, hypothesis testing risks two fundamental errors. First, researchers can conclude that two treatments differ when, in fact, they do not. This **type I error** (α) measures the probability of making this false-positive conclusion. Conventionally, α is most frequently set at 0.05, meaning that investigators desire a less than

5% chance of making a false-positive conclusion. Second, researchers can conclude that two treatments do not differ when, in fact, they do—that is, a false-negative conclusion. This **type II error** (β) measures the probability of this false-negative conclusion. Conventionally, investigators set β at 0.20, meaning that they desire less than a 20% chance of making a false-negative conclusion (Schulz and Grimes 2005).

Power derives from β error. Mathematically, it is the complement of β ($1 - \beta$) and represents the probability of avoiding a false-negative conclusion. For example, for $\beta = 0.20$, the power would be 0.80, or 80%. Stated alternatively, power represents the likelihood of detecting a difference (as significant, with $P < \alpha$), assuming a difference of a given magnitude exists. For example, a trial with a power of 80% has an 80% chance of detecting a difference between two treatments if a real difference of assumed magnitude exists in the population (Schulz and Grimes 2005).

Table 4.11 Some strategies for decreasing sample size required in a clinical trial (Scales and Rubenfeld 2005) (Reprinted with permission from Elsevier)

Strategies for decreasing sample size	Major drawbacks
Increase the minimal clinically important difference	Raises likelihood of study being negative
Increase acceptable type I error rate	Increases likelihood of incorrectly assuming treatment effect despite none existing
Increase acceptable type II error rate	Increases likelihood of incorrectly assuming treatment is ineffective
Reduce SD • Conduct paired comparisons • Use a more homogeneous population • Use a more precise measurement • Take multiple measurements of the primary end point and average the results	• Can decrease generalizability • Recruitment may be more difficult; can decrease generalizability • May increase cost of study • May increase cost of study
Choose an outcome for which a greater difference between groups is expected • Use physiological variable • Use surrogate markers • Use composite end points	End point may no longer be ‘clinically relevant’
Replace a dichotomous variable with a continuous one	• The continuous end point may be less convincing to clinicians • SE of measuring the primary end point must be determined

Table 4.12 Summary of the components for sample size calculations (Noordzij et al. 2010) (Reprinted with permission from Oxford University Press)

Component	Definition
Alpha (type I error)	The probability of falsely rejecting H_0 and detecting a statistically significant difference when the groups in reality are not different, i.e. the chance of a false-positive result.
Beta (type II error)	The probability of falsely accepting H_0 and not detecting a statistically significant difference when a specified difference between the groups in reality exists, i.e. the chance of a false-negative result.
Power (1-beta)	The probability of correctly rejecting H_0 and detecting a statistically significant difference when a specified difference between the groups in reality exists.
Minimal clinically relevant difference	The minimal difference between the groups that the investigator considers biologically plausible and clinically relevant.
Variance	The variability of the outcome measure, expressed as the SD in case of a continuous outcome.

Abbreviations: H_0 null hypothesis; i.e. the compared samples come from the same source population (the compared groups are not different from each other), SD standard deviation

4.1.7.2 Components of Sample Size Calculation

The minimum information needed to calculate sample size for a randomized controlled trial in which a specific event is being counted includes (Endacott & Botti 2007; Noordzij et al. 2010, 2011) (Table 4.12):

1. **The significance level** – normally set at 0.05 or 0.01 (P value). Statistical significance refers to reducing the probability of finding an effect when the effect does not exist, referred to as type I error. When the level of significance is 0.05, the probability of a type I error is 5%.

Table 4.13 Summary of key points for sample size calculation (Kelly et al. 2010) (Reprinted with permission from John Wiley & Sons, Inc.)

• Sample size calculation should reflect the analysis that will be conducted
• Power should be at least 80%
• Treatment difference to be detected should be based on a difference that is clinically relevant, not on what is expected
• Attempt to obtain good estimates of the information that is required for a sample size calculation (e.g. standard deviation, control response rate)
• Try to be conservative
• Adjust for dropouts
• Be clear and explicit about how your sample size was calculated

2. **Power** – normally set at greater than 0.8 (80%). Power refers to the probability of finding an effect when the effect does exist. In situations of low power (inadequate sample size), important effects may not be detected. This is referred to as type II error.

3. **Effect size refers to the minimum size of the difference** to be detected and is determined by the clinical importance of the difference. The effect size may be derived from a similar study, pilot work, or estimates of the likely effect.

Planning sample size can be accomplished through (at least) two conceptually different methods, one designed to obtain statistical power and the other designed to obtain statistical precision. Depending on the particular theoretical question of interest and the desired goals of a study, sample size planning should be approached from the power analytic approach, the AIPE approach, or a combination of the two. It is important to realize that planning sample size from one of the approaches is a fundamentally different task than planning from the other. The distinction between the two approaches is more than conceptual, as the differences in estimated sample sizes can be substantial depending on the desired level of power and the desired width of the confidence interval (Kelley et al. 2003).

The calculated sample size should then be adjusted for other factors, including expected compliance rates and less commonly, an unequal allocation ratio (Kirby et al. 2002).

4.1.7.3 Allocation Ratio

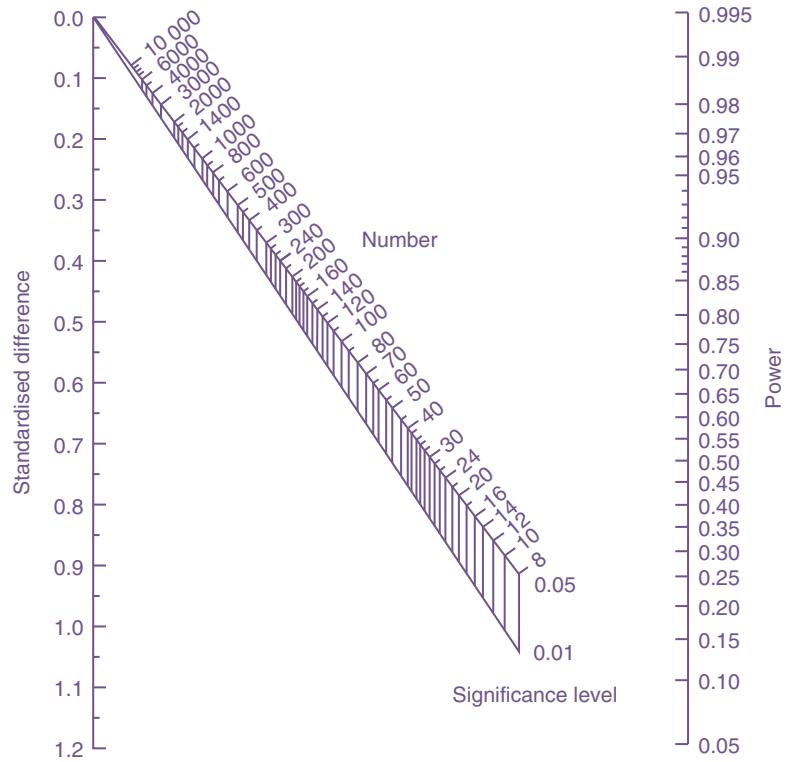
The ratio of the number of patients in one group compared to the other is called the allocation

ratio. The most efficient allocation of patients is equal numbers in each group (a ratio of 1:1), as this will require the smallest required total sample size. However, sometimes it might be beneficial to have more patients in one treatment arm (Kelly et al. 2010). This may be justified where the investigational treatment is unusually expensive or complicated to administer (Kirby et al. 2002). Commonly used allocation ratios are 1:1, 2:1 and 3:1. Using a ratio greater than 4:1 is not advised: There will be only a small decrease in the number of patients required in the smaller control arm at the expense of a very large increase in the overall total sample size (Kelly et al. 2010).

4.1.7.4 Methodology for Sample Size Calculation

In RCTs, a lot of money is invested, and it is therefore important to be sufficiently sure that enough patients are included in the study arms in order to find as statistically significant difference that we assume there is in the population (Florey 1993; Noordzij et al. 2010, 2011). Sample size calculations should be performed during the design phase of the study. There are several computerized methods that can be easily accessed through the Internet if statistical support is not at hand. However, the formula must be appropriate for the design of the trial and the subsequent analysis. Each sample size calculation is unique, and there is no global solution applicable to all trials (Karlsson et al. 2003). In general, sample size calculations are performed based on the primary outcome of the study (Noordzij et al. 2010, 2011). **Table 4.13**

Fig. 4.5 Nomogram for the calculation of sample size or power (Altman 1980. Reprinted with permission from BMJ Publishing Group Ltd.)



summarizes some of the key points for sample size calculation (Kelly et al. 2010).

It is strongly recommended that one enlist the aid of a biostatistician at the beginning to help determine the correct sample size calculations. However, it is always a good idea to understand the calculation and why the statistician did what he or she did. Things to consider when determining sample size (Karlsson et al. 2003):

1. The investigation or analytical procedure that one wishes to follow as part of the initial experimental design stage of the investigation
2. The statistical procedure that one may wish to use for analysis (e.g., 1-tailed or 2-tailed *t*-test, linear regression)
3. The percentage change and/or level of significance at which one would wish to detect the change
4. The appropriate sample size formula for the study design in question

For some simple clinical trials, **nomograms** or graphs can be used to estimate the sample size required for the study. An example of such nomo-

gram published by Altman (1980) is presented in **Fig. 4.5**. However, one should keep in mind that, although these graphical methods work well, they often make assumptions about the type of data and statistical tests to be used (Noordzij et al. 2010). To use this nomogram, one needs the standardized difference, which can simply be calculated by dividing the minimal clinically relevant difference (0.50 g/dl) by the SD in the population (1.90 g/dl). For our example, this yields $0.50/1.90=0.26$. We can now use the nomogram to estimate the sample size by drawing a straight line between the value of 0.26 on the scale for the standardized difference and the value of 0.80 on the scale for power and reading off the value on the line corresponding to $\alpha=0.05$, which gives a total sample size of 450, that is, 225 per group. However, although this nomogram seems to work well for this example, one should keep in mind that these graphical methods often make assumptions about the type of data and statistical tests to be used. In many cases, it is therefore more appropriate to apply statistical formulas to calculate the required sample size (Noordzij et al. 2011).

Several software programs such as nQuery Advisor and PASS can assist in sample size calculations for different types of data and study designs. In addition, there are some websites that allow free sample size calculations, but not all of these programs are reliable. However, because many methods are not straightforward, we recommend consulting a statistician in all but the most basic studies (Noordzij et al. 2011).

Finally, one should explicitly choose between 2-sided or 1-sided statistical testing. Unfortunately, the importance of this decision is often neglected. In the 2-sided option, one is able to detect the predefined relevant difference, if present, in both directions. That is, if the intervention is better than the reference of a given amount, this will probably be found, but also a similar advantage of the reference over the principal intervention can be detected. The 1-sided option would only allow a 1-sided evaluation (e.g., whether or not the intervention is better than the reference treatment of a given amount), without testing whether the reference is superior to the principal intervention (Karlsson et al. 2003).

For dental studies, the sample size needs to be specified at two levels: (1) the number of patients (n) enrolled in the study and (2) the average number of sites (m) examined per patient. In general, m and n should be selected in such a way that the precision of the research findings is maximized, while the cost of the study is minimized. When all sites of interest are measured within the sampled patients, the sampling model corresponds to a one-stage cluster design (whole-mouth recordings). Selecting all sites within a patient may not be cost-effective, in which case a sample of sites within each patient should be taken. Such a design is called a two-stage sample or a partial-recording procedure. For such designs, the patients are called the first-stage units and the sites the second-stage units. A first question which arises in the design of dental studies is: ‘Should all sites (teeth) within a patient be sampled, or should sites (teeth) be subsampled?’ This question may be answered by determining the average number of sites that should be examined per patient to maximize the precision for a fixed cost. This number will be denoted as the optimum site sample size. When the optimum site sample size is approximately equal to or in excess of the average number

of available sites per patient, then whole-mouth examinations should be performed (resulting in a one-stage cluster sample). However, when the optimum site sample size is significantly lower than the average number of available sites per patient, then partial-recording procedures should be performed (corresponding to a two-stage cluster sample) (Hujoel and DeRouen 1992).

4.1.7.5 Challenges in Sample Size Calculation

There are several difficulties in obtaining an appropriate and realistic sample size (Kelly et al. 2010):

1. **Choosing a value for treatment difference:** Ideally, the difference to be detected should be the minimum difference that is clinically important. The larger the difference used for the sample size calculation, the greater the chance that the trial will not find a significant result, when in fact there is a difference (false-negative result).
2. **Choosing values for the additional information required:** One of the biggest problems with sample size calculations is that we are required to know the variability in the data in advance of the study being conducted. Knowing the standard deviation or the proportion of patients that have the outcome of interest prior to the study is clearly impossible. This can be done by considering data that already exists, for instance, from results of similar previous studies (e.g., published work) or from routinely collected health data, such as hospital or registry data. Another option is to conduct a small pilot study to generate some estimates
3. **Estimating recruitment rate:** The recruitment rate will define how long the trial will need to run to recruit the required number of participants. Usually, recruitment is slower at the beginning of a study and speeds up as the trial gains momentum.
4. **Withdrawals:** Withdrawals are patients who for known or unknown reasons are lost to follow up, and so the endpoint of interest is unknown for that patient. Withdrawal usually occurs either due to lost contact (e.g., if the patient has moved location) or after an intervening event prevents gathering the necessary information.

5. **Non-adherence (compliance):** Unlike withdrawals, non-compliance refers to patients remaining in the study such that their outcome is measured, but he/she has not adhered to the treatment protocol, for example, by not taking the drug as specified or by switching therapies. Again, the sample size can be adjusted accordingly.
6. **Choosing the primary endpoint:** Investigators are often interested in more than one outcome. In some cases, if there are several outcomes equally important, a sample size can be calculated for each outcome and the largest sample size chosen.
7. **Target sample size is not achieved:** Sometimes it is difficult to achieve the target sample size due to recruitment problems. As a result, the study will be underpowered (unless some of the sample size assumptions were incorrect).

Study Sample Size: Examples

- *The primary outcome was clinical attachment level (CAL). Study sample size calculations were based on subject-level analyses (t-tests of aggregate mean scores at a single follow-up) as the study randomized individuals. Under the alternative hypothesis, a mean difference in the observed CAL between groups of 1 mm, with a standard deviation (SD) of 1 mm, would require 16 or 22 individuals in each group to detect a significant difference at the 5% level with 80% or 90% power, respectively. With multiple follow-up measures and the intention-to-analyse sites within a multilevel framework, study-size estimates based on subject-level t-tests were anticipated to be overly conservative. Consequently, while the study sought to recruit up to 22 individuals per group, achieving 16 and 18 in the treatment and control groups, respectively, was deemed satisfactory (Needleman et al. 2007).*
- *Sample size calculation took into consideration estimated alpha error of 5% and beta error of 20% to detect a post-therapy difference of 1.0 mm, between treatment groups, in the means of clinical attachment level (CAL) at sites exhibiting baseline CAL of 7 or more millimeters, for an expected within-group standard deviation of 1.0 mm. For a two-sided test, 17 subjects were needed in each group and a total of 40 were to be selected to compensate*

for a possible 15% drop out rate during the course of the study (Varela et al. 2011).

- *A power calculation, with data from a previous study, based on the detection of a difference in the mean gingival index interproximally of 20% between treatment groups indicated that 75 participants were required in each group ($\alpha = 0.05$, $\beta = 0.2$) (Jönsson et al. 2010).*

4.1.7.6 Sample Size Calculations for Special Types of RCTs

Alternative approach for sample size calculations are needed for special types of RCTs, such as (Noordzij et al. 2010):

- **Cluster-randomized trials**, in which health interventions are allocated randomly to intact clusters, communities, health centres, or practices rather than to individual subjects (Hayes and Bennett 1999).
- **Crossover design**, because these studies compare the results of two treatments on the same group of patients. The sample size calculated for a crossover study can also be used for a study that compares the value of a variable after treatment with its value before treatment (Giraudeau et al. 2008).
- **Clinical trials testing the equivalence rather than the superiority** between two treatments need another approach. These equivalence or non-inferiority trials usually demand higher sample sizes (Christensen 2007).

The aforementioned alternative calculations are less common and more complicated and will in most cases require statistical assistance (Noordzij et al. 2010).

4.1.7.7 The Number Needed to Treat (NNT)

The effect of a new treatment, relative to that of the standard of care or a placebo control, can be presented in many different ways using a variety of summary statistics. For example, in periodontal research, the effect of an intervention relative to that of a control may be based on differences between the study groups in the proportion of sites with ≥ 2 mm attachment loss (risk difference), ratios of the proportions of sites with ≥ 2 mm attachment loss (relative risk), or differences in mean attachment loss. Another measure of treatment effect, used to

summarize the clinical benefit of a treatment, is the number needed to treat (NNT) (Stoner and Payne 2008; Chatellier et al. 1996). The NNT represents the number of persons (number of sites) that must be treated during a given period to achieve an event (treatment results) or to prevent an adverse effect (Greenstein and Nunn 2004). This topic is further developed in **Chap. 9: Statistical versus Clinical Significance in Periodontal Research**.

4.1.8 Sample Size Changes in RCTs

Randomized controlled trials commonly have two types of exclusions, those before and those after randomization. The impact of these types of exclusions differs (Grimes and Schulz 2002).

Exclusions before randomization are common in most randomized controlled trials. Volunteers are admitted to randomized controlled trials only after having passed eligibility criteria specified in the study protocol. These criteria range from medical (e.g., one treatment would not be safe for the potential participant) to logistical (e.g., the prospective participant is deemed unlikely to return for follow-up visits). Whether capricious or reasonable, none of these exclusion criteria will bias the study: The internal validity should be intact if the trial is properly done. What may suffer from extensive exclusions before randomization is external validity: The trial may include such an eclectic sample that readers cannot extrapolate the results to their patients (Grimes and Schulz 2002).

In contrast, **exclusions after randomization** open the door to bias. Many investigators and readers fail to appreciate that the only unbiased comparison groups are those initially composed by the randomization. Any attrition after that leads to groups that are no longer comparable (unless the erosion is random, which is usually not the case) (Schulz and Grimes 2002d).

Several common causes of exclusions after randomization are (Schulz and Grimes 2002d; Grimes and Schulz 2002):

- **Eventual discovery of ineligibility.** In some trials, participants are enrolled and later discovered not to have met the eligibility criteria. Exclusions at this point could seriously bias the results since discovery is probably not random.

For example, participants least responsive to treatment or who have side effects might draw more attention and therefore, might be more likely to be judged ineligible than other study participants. Alternatively, a physician who had treatment preferences for certain participants might withdraw individuals from the trial if they were randomly assigned to what he believes to be the wrong group.

- **Protocol deviations.** Deviations from assigned treatment happen in many trials. Some investigators suggest that participants who deviate substantially from the allotted treatment should be excluded in the final analysis or should be included only up to the point of deviation. Although this approach seems attractive, it has a serious flaw: ‘The group which deviates from one protocol and the group which deviates from the other protocol may be so different [...] that the treatment comparison in the remaining patients will be severely biased’.
- **Losses to follow-up.** Participants might move or might refuse to continue participating in the trial. Without outcomes from those lost to follow-up, investigators have little choice but to exclude them from the analysis. Any losses damage internal validity, but differential rates of loss among comparison groups cause major damage. Many investigators could work harder than they do to obtain higher follow-up rates (Table 4.14). How much loss to follow up is too much? Some have suggested a *5-and-20* rule of thumb. Trials with less than 5% loss to follow up are probably secure; those with rates higher than 20% may be unsalvageable.

Trialists should endeavour to minimize exclusions after randomization and to do intention-to-treat analyses. They should also follow the CONSORT Statement for reporting. The flow diagram (trial profile) helps particularly to track the progress of participants through a trial (Schulz and Grimes 2002d).

According to the CONSORT glossary, **intention-to-treat analysis (ITT)** is a strategy for analysing data in which all participants are included in the group to which they were assigned, whether or not they completed the intervention given to the group. Intention-to-treat analysis prevents bias caused by the loss of

Table 4.14 Approaches to maximisation of participant follow-up (Schulz and Grimes 2002d) (Reprinted from The Lancet 2002;359(9308):781–5 with permission from Elsevier)

• Hire a person to manage and encourage follow-up
• Hire personnel to call participants or to visit participants at their homes or place of work, if participants are not returning for follow-up
• Exclude before randomisation those likely to be unwilling to return
• Exclude before randomisation those likely to move
• Obtain contact information to prompt participants to return for follow-up and to facilitate location of participants if they do not return – e.g., mail, telephone, and e-mail for enrolled participants, for close friends or relatives who do not live with the participant, and for the participant’s family doctor
• Obtain an identification number, such as a national health-care number
• Establish follow-up venues suited to participants rather than to investigators and trial implementers – e.g., more locations than just the central clinic or hospital, close to where participants live, convenient to access, and sensitive to waiting time
• Streamline trial procedures to move participants quickly through a follow-up visit
• Keep data collection instrument short so as to not overburden the participant
• Provide excellent and free medical care
• Provide monetary subsidies, primarily for time and travel costs incurred by participants

participants, which may disrupt the baseline equivalence established by random assignment and which may reflect non-adherence to the protocol (Heritier et al. 2003). ITT analyses are generally preferred as they are unbiased and also because they address a more pragmatic and clinically relevant question. An analysis in which data are analysed for every participant for whom the outcome was obtained is often described as an **available case analysis**. Some trial reports present analyses of the results of only those participants who completed the trial *and* who complied with (or received some of) their allocated intervention. Some authors incorrectly call this an ITT analysis, but it is in fact a **per-protocol analysis**. Furthermore, some authors analyse participants only according to the actual interventions received, irrespective of the randomized allocations (**treatment-received analysis**). It is generally unwise to accept study authors’ description of an analysis as ITT; such a judgement should be based on the detailed information provided (Cochrane Handbook 2011).

Three arguments against intention-to-treat analysis were made by Van der Weijden et al. (2008): In a trial intended for ‘proof of principle’, ITT may obscure the relevant test; ITT may give investigators an incentive to coerce subjects; and because adherence in a study may differ from adherence in prac-

tice, ITT may give a misleading result. As a comment addressed to these arguments, Hodges et al. (2008) demonstrated in two examples that ITT was violated for the sake of proving a principle, in both cases giving the wrong result in terms of patient outcomes. It was suggested that investigators should have ITT in mind from the first moment of study planning. Doing so might reduce or even eliminate the apparent contradiction between ITT and subject autonomy (Hodges et al. 2008).

Participant flow diagrams are an effective method of summarizing all stages of key trial processes. They suggest how these processes should be reported in the trial, specifying separately trial enrolment, treatment allocation, subject follow-up, and statistical analysis. In reporting the results of randomized studies, all individuals originally considered for participation in the study should be accounted for in the participant flow diagram. The diagram also provides an overview of many aspects of study quality that can have a major influence on the generalizability and reliability of the conclusions (Cakir et al. 2003).

When the quality of RCTs in periodontology was assessed, using evidence-based components, it was revealed that approximately half of the studies (56%) demonstrated proper accounting for all participants at the end of the trial. For the rest, accounting of participants was absent in 14% of

studies and impossible to determine for 30% of trials. The use of analyses to take into account withdrawals (dropouts, losses to follow-up, excluded patients) was not applicable for 43% of studies where follow-up was deemed complete. For the rest, only 11% of analyses took withdrawals into account, although a larger proportion (56%) provided inadequate information to determine whether this had been done (Montenegro et al. 2002).

4.2 Clinical Trial Implementation

4.2.1 Manual of Operations

One of the first things faced in the implementation of a clinical trial is development of a Manual of Operations (sometimes referred to as a Manual of Procedures or MOP). This essential document serves as a 'cookbook' for the trial and describes the procedures to be used in the trial. It is a guide to trial procedures and is a reference for all investigators and staff involved in the trial. In addition to the trial background and protocol, it includes information such as site locations and facilities, policies concerning trial committees, advisory boards, IRB approvals, costs, patient inclusion/exclusion criteria, treatment procedures, evaluation and follow-up, data management and analysis, medication handling and storage, data collection forms, and publication policy (Pihlstrom and Barnett 2010).

A description of components of a Manual of Operations is given on websites by the National Institute of Allergy and Infectious Diseases, National Institutes of Health Bethesda, Maryland (Manual of Operations For Therapeutic Clinical Trials: <http://www.knepk.litbang.depkes.go.id/knepk/kegiatan/basic/templates/manual.pdf>), the Case Comprehensive Cancer Center (Clinical Trials Operations Manual: <http://cancer.case.edu/researchadmin/redbook.pdf>), or the Obstetrics and Periodontal Therapy (OPT) Study (<http://www.biostat.umn.edu/~hodges/MOP-IndexVersion07-08-03.doc>).

Patient Selection: Examples

- *The inclusion criteria were as follows: age 30–70 years, at least 16 teeth present, diagnosis of chronic periodontitis, at least two teeth with*

probing depth (PD) of ≥ 6 mm and at least 30% bone loss in at least two quadrants excluding third molars, current smoker of at least 10 cigarettes per day for a minimum of 1 year. Exclusion criteria were: allergy to tetracycline, antibiotic prophylaxis required before periodontal therapy, daily consumption of non-steroidal anti-inflammatory drugs or antibiotics and pregnancy or lactation and any health condition potentially affecting the response to periodontal therapy (e.g. diabetes mellitus) (Needleman et al. 2007).

- *Included patients were diagnosed with generalized aggressive periodontitis according to criteria of the American Academy of Periodontology. Further inclusion criteria were 1) age between 18 and 39; 2) presence of at least 16 teeth and 4 sites on different teeth (three of them other than central incisors or first molars) with a probing pocket depth ≥ 6 mm and clinical attachment level ≥ 5 mm, bleeding on probing (BOP). The exclusion criteria were: 1) reported allergy to penicillin, metronidazole (MET) or chlorhexidine (CHX); 2) systemic conditions that could modify the progression or treatment of periodontal diseases, including diabetes and immunodeficiencies; 3) need of antibiotic coverage for periodontal procedures; 4) long term use of anti-inflammatory medication; 5) periodontal treatment and/or use of antibiotics in the last 6 months; 5) pregnancy; and 6) breastfeeding (Varela et al. 2011).*

4.2.2 Clinical Trial Registration

Publication bias is the selective publishing of favourable or statistically significant results. This practice, over time, distorts the medical literature by depicting inordinately optimistic outcomes for treatments and interventions. Sources of publication bias include preferential publishing by journals and preferential submission by researchers (Abaid et al. 2007).

Mandatory trial registration was instituted by the International Committee of Medical Journal Editors (ICMJE), to reduce publication bias by improving the ability to identify all trials pertaining to a specific intervention. Trials must register at or before the onset of patient enrolment. This policy applies to any clinical trial starting

enrolment after July 1, 2005. For trials that began enrolment before this date, the ICMJE member journals will require registration by Sept. 13, 2005, before considering the trial for publication (De Angelis et al. 2004a, b, c).

The ICMJE does not advocate one particular registry, but its member journals will require authors to register their trial in a registry that meets several criteria. The registry must be accessible to the public at no charge. It must be open to all prospective registrants and managed by a not-for-profit organization. There must be a mechanism to ensure the validity of the registration data, and the registry should be electronically searchable. An acceptable registry must include at minimum the following information: a unique identifying number, a statement of the intervention (or interventions) and comparison (or comparisons) studied, a statement of the study hypothesis, definitions of the primary and secondary outcome measures, eligibility criteria, key trial dates (registration date, anticipated or actual start date, anticipated or actual date of last follow-up, planned or actual date of closure to data entry, and date trial data considered complete), target number of subjects, funding source, and contact information for the principal investigator (De Angelis et al. 2004a, b, c).

These key data requirements were superseded by the minimal data registration set released in April 2005 by the World Health Organization (WHO). There are currently 20 items in the **WHO Trial Registration Data Set**. It is sometimes referred to as the **TRDS**:

1. Primary Registry and Trial Identifying Number
2. Date of Registration in Primary Registry
3. Secondary Identifying Numbers
4. Source(s) of Monetary or Material Support
5. Primary Sponsor
6. Secondary Sponsor(s)
7. Contact for Public Queries
8. Contact for Scientific Queries
9. Public Title
10. Scientific Title
11. Countries of Recruitment
12. Health Condition(s) or Problem(s) Studied
13. Intervention(s)
14. Key Inclusion and Exclusion Criteria
15. Study Type

16. Date of First Enrollment
17. Target Sample Size
18. Recruitment Status
19. Primary Outcome(s)
20. Key Secondary Outcomes

Currently, there are three types of trial registries:

1. Government registries such as that maintained by the US NIH in collaboration with the US FDA (ClinicalTrials.gov 2011)
2. Registries developed by publishing companies (e.g., Current Controlled Clinical Trials 2011) and
3. Registries maintained by pharmaceutical companies (e.g., GlaxoSmithKline. Clinical Trial Registry 2011) (Pihlstrom and Barnett 2010).

The World Health Organization (2011) has taken on the task of creating an international system of linked registries, recognizing that multiple fragmented registries would do little to overcome publication bias (International Clinical Trials Registry Platform 2011) to provide information about clinical trials (Abaid et al. 2007).

Clinical Trial Registration: Examples

- *The study was prospectively registered on the Current Controlled Trials Register in October 2002 (ISRCTN11033714) (Needleman et al. 2007).*
- *This study was approved by the University of Pennsylvania Institutional Review Board, and is registered at <http://clinicaltrials.gov> (registration number NCT00116974, dated 30 June 2005) (Jeffcoat et al. 2011).*

4.2.3 Ethical Standards and Informed Consent

Numerous ethical standards must be met in the conduct of a clinical trial; all staff must be fully acquainted with procedures and requirements to protect those who volunteer to participate in clinical trials. The basis for human participant protections can be found in the Nuremberg Code of 1947, a 10-point statement defining conditions under which medical experiments using humans would be permissible (Nuremberg Code, 1996) (Pihlstrom and Barnett 2010). All countries con-

duct device clinical investigations in accordance with the Helsinki Declaration adopted by the 18th World Medical Assembly in Helsinki, Finland, in 1964. This Declaration was last amended by the 41st World Medical Assembly in Hong Kong in 1989 (Cooley and Castellion 1997).

Good Clinical Practice (GCP) is an international quality standard that is provided by International Conference on Harmonization (ICH). This is an international body that defines standards on how clinical trials should be conducted, which governments can transpose into regulations for clinical trials involving human subjects. These guidelines include protection of human rights as a subject in clinical trials. It also provides assurance of the safety and efficacy of the newly developed compounds. One of the basic aims of GCP is to protect the subjects involved in research (Van der Weijden et al. 2008).

4.2.3.1 Informed Consent

Informed consent is necessary to ensure that all participants understand the study, its potential risks and benefits, and that they can withdraw without penalty from the study at any time (Pihlstrom and Barnett 2010). A number of different issues should be addressed during the informed consent process. Some institutional review boards suggest that these elements be covered in the consent form using an outline format with headings to clearly identify the purpose of each element (Dennison 1997):

- Invitation to participate in a research study
- Purpose of the study
- Alternative treatments
- Description of their involvement in the study
- Time commitment
- Mechanisms for orderly withdrawal from the study
- Termination of participation
- Benefits
- Risks
- Compensation of injury
- Financial considerations
- Confidentiality/anonymity
- Contact person for further information about the study
- Final statement

4.2.3.2 Consent to Randomization

As a general rule, prospective subjects of RCTs should be informed that their therapy will be chosen by chance. The primary reason for this is that, although the therapies to be compared are considered medically equivalent by investigators, the composition of risks and benefits usually differs in ways such that reasonable prospective subjects might choose one or another based upon their personal values. Moreover, patients expect that physicians will provide advice based upon their personal knowledge of the patient, his or her ailment, and of therapeutics. If the state of relevant knowledge in therapeutics is such that choice of therapy by chance is justified, this fact, of itself, should be disclosed.

4.2.3.3 The Nature of Controls

The Declaration of Helsinki, first drawn up in 1964, has become a standard guidance for all medical research that involves human beings. In particular, it has been widely used to guide the conduct of clinical trials done to develop new medicinal products. The revision of the declaration, issued by the World Medical Association in October 2000, has been appropriately strengthened in several areas, and its relevance and value remains as great as ever. However, the new version contains one section, section 29, which could cause substantial difficulties for the future development of medical products if it is interpreted literally and implemented universally. Section 29 is in the part of the declaration that deals with principles for medical research combined with medical care (Lewis et al. 2002):

The benefits, risks, burdens, and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods. This does not exclude the use of placebo, or no treatment, in studies where no proven prophylactic, diagnostic, or therapeutic method exists.

Although it allows the testing of new treatments, Article 29 perpetuates the proscription of placebo controls in clinical trials when even partially effective treatments exist. Reacting to expressions of dismay over Article 29, in 2001, the WMA issued a clarification regarding its

interpretation. The clarification reads as follows: The WMA hereby reaffirms its position that extreme care must be taken in making use of a placebo-controlled trial and that in general this methodology should only be used in the absence of existing proven therapy (Carpenter et al. 2003). However, a placebo-controlled trial may be ethically acceptable, even if proven therapy is available, under the following circumstances:

Where for compelling and scientifically sound methodological reasons its use is necessary to determine the efficacy or safety of a prophylactic, diagnostic or therapeutic method; or Where a prophylactic, diagnostic or therapeutic method is being investigated for a minor condition and the patients who receive placebo will not be subject to any additional risk of serious or irreversible harm.

All other provisions of the Declaration of Helsinki must be adhered to, especially the need for appropriate ethical and scientific review (WMA website: <http://www.wma.net/e/home.html>) (Carpenter et al. 2003).

This clarification narrows the injunction against placebo substitution for a proven treatment. Clinical trials with a placebo control group would be ethically justified where there are good reasons for placebo use or if the condition being studied is ‘minor’ and the additional risk is negligible. This is an important recognition of a moral distinction among circumstances in which placebos may be used. It would require a case-by-case review of the justifications for placebo use, a process sensitive to scientific merit and study-specific risk based on subject selection and risk-reduction procedures (Carpenter et al. 2003).

Levine et al. (2003) preferred the much clearer guidance on this matter provided in the Council of International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines for Biomedical Research Involving Human Subjects (revised version issued in 2002). The relevant passage is the following:

An exception to the general rule (requiring established effective treatment as the control condition) is applicable in some studies designed to develop a therapeutic...intervention for use in a country... in which an established effective intervention is not available and unlikely in the foreseeable future to become available, usually for economic or logistic reasons. The purpose of such a study is to make

available to the population of the country... an effective alternative to an established effective intervention that is locally unavailable. Accordingly, the proposed investigational intervention must be responsive to the health needs of the population from which the research subjects are recruited and there must be assurance that, if it proves to be safe and effective, it will be made reasonably available to that population. Also, the scientific and ethical review committees must be satisfied that the established effective intervention cannot be used as comparator because its use would not yield scientifically reliable results that would be relevant to the health needs of the study population. In these circumstances an ethical review committee can approve a clinical trial in which the comparator is other than an established effective intervention, such as placebo or no treatment or a local remedy.

The guidelines issued by the European Group are generally compatible with, albeit not as specific as, the CIOMS statement. Each of these ethics statements now seems to require a protocol-by-protocol judgement of the ethical issues, rather than a categorical rejection of the use of placebo where standard treatment exists (Levine et al. 2003).

Several institutions opted to cite earlier versions of the Declaration, and the US Food and Drug Administration (FDA) recently removed all reference to the declaration in its approval requirements for drugs and biological products that are studied outside the United States. These developments suggest a significant weakening of the declaration, which is concerning because strong international guidance on the ethics of medical research is sorely needed. However, the recent 2008 revision makes a problematic attempt to strengthen the declaration. By asserting that ‘no national or international ethical, legal or regulatory requirement should reduce or eliminate any of the protections for research subjects set forth in this Declaration’ (§10), the declaration seems to position itself as ‘first among equals’ in research ethics and regulation (Table 4.15) (Rid and Schmidt 2010).

There are well-known reasons why placebo-controlled trials are generally more reliable than active-controlled trials in provision of conclusive proof of efficacy. The reasons are fully discussed in the E10 guideline *Choice of Control Group*, which was produced under the auspices of the International Conference on Harmonization (ICH) and has been adopted in Europe, USA, and

Table 4.15 Selected changes in the 2008. Declaration of Helsinki (DoH) for better orientation, the changes have been organized according to four major themes. However, there is significant overlap between categories (Rid and Schmidt 2010) (Reprinted with permission from John Wiley & Sons, Inc.)

Changes related to ...	Topic	DoH 2004 1	DoH 2008
1. Scope and status of DoH	Audience	Physicians and other participants in medical research (§1)	Primarily physicians (§2)
	Authority	Primacy over national ethical, legal or regulatory requirements (§9)	Primacy over national and international ethical, legal or regulatory requirements (§10/15)
2. Access to research and results	Access to research	–	Ensure appropriate access of underrepresented populations (§5)
	Dissemination of research results	–	Publicize research results (§30)
	Information about study outcomes	–	Inform participants about study outcomes (§33)
3. Subject protection	Balance of individual and societal interests	Well-being of individual participant overrides scientific and societal interests (§5)	Well-being of individual participant overrides ‘all other interests’ (§6)
	Vulnerable populations	Special protection for participants who are economically or medically disadvantaged, unable to consent, consent under duress, will not benefit from research, engage in research combined with medical care (§8)	Special protection for participants who are unable to consent or vulnerable to coercion or undue influence (§9)
	Compensation for research injuries	–	Make provisions for treating/compensating for research injuries (§14)
	Review of protocol changes	–	Ethical review of all protocol changes (§15)
	Trial registration	Publicize study design (§16)	Register trial (§19)
	Exemptions from consent (competent subjects)	–	Only if (1) consent is impossible/impractical to obtain or undermines scientific validity and (2) ethical review has occurred (§25; research with identifiable human material/data)
	Research with incompetent subjects	Only if (1) necessary to promote health of the population and (2) research cannot be done with competent persons (§24; all research with incompetent subjects)	Only if (1) necessary to promote health of the population and (2) research cannot be done with competent persons and (3) no more than minimal risk and burden (§27; research with incompetent subjects that offers no prospect of direct benefit)
	Research with incompetent subjects	Obtain assent and surrogate consent (§25)	Obtain assent, respect dissent and obtain surrogate consent (§28)
	Research combined with medical care	Only if research is justified by its potential prophylactic, diagnostic or therapeutic value for subject (§28)	Only if (1) research is justified by its potential preventive, diagnostic or therapeutic value for subject and (2) participation will not adversely affect subject’s health (§31)
	Study withdrawal	–	Must never interfere with patient-physician relationship (§34)

(continued)

Table 4.15 (continued)

Changes related to ...	Topic	DoH 2004 1	DoH 2008
4. International research collaborations	Research and benefit for the population	Research only if population benefits from research results (§19; <i>all research</i>)	Research only if (1) responsive to health needs and priorities of the population and (2) reasonable likelihood that population will benefit from research results (§17; <i>research involving disadvantaged or vulnerable populations</i>)
	Family/community consent	–	Where appropriate, family/ community consent can be obtained, but does not substitute or override individual consent (§22)
	Placebo-controlled trials	Only if (1) no proven prophylactic, diagnostic or therapeutic method or (2) compelling scientific reasons or (3) no additional risk of serious or irreversible harm (§30 with note of clarification)	Only if (1) no proven prophylactic, diagnostic or therapeutic method or (2) compelling scientific reasons and no additional risk of serious or irreversible harm (§32)
	Benefit sharing/ post-trial access	Obligation to ensure subjects access to the best proven prophylactic, diagnostic and therapeutic methods identified by the study (§30)	Entitlement to share any benefits resulting from research, for example, access to interventions identified as beneficial in the study or other appropriate care or benefits (§33)

Japan. Briefly, placebo-controlled trials provide their own check of internal validity. A positive difference in efficacy, when detected in a well-designed and implemented placebo-controlled trial, means, first, that the trial is capable of detecting differences, and second, that the test treatment is more efficacious than placebo. Furthermore, for regulatory purposes, a satisfactory analysis of a placebo-controlled trial will be judged conservative—that is, to indicate the minimum effect plausibly due to the test treatment. This judgement ensures that treatments with statistically significant but clinically inadequate effects are identified and results in appropriately cautious licensing decisions (Lewis et al. 2002).

In a clinical trial, approval must be obtained from the Institutional Ethics Committee (IEC) at the individual clinical centres. Each committee must approve the protocol and its own informed consent document, which may have different, culturally sensitive, language for its participant population. Coordinating the requirements of multiple ethics committees can be a major challenge because some trials may have many clinical

centres or data collection sites (Pihlstrom and Barnett 2010).

In periodontology, at first glance, defining controls for a clinical study would appear simple (Meinert 1985). One group is randomly assigned to receive test therapy and another group receives the control therapy. In fact, the definition of controls for stand-alone mechanical therapy is difficult. Given the long-standing history of the efficacy of scaling and root planning (SRP), the ethics of randomized control with a no-therapy group is an important issue. Furthermore, in the case of mechanical or surgical treatments, the experienced examiner can usually see which patients have received therapy. Under these circumstances, where regulatory approval is not required, many studies utilize historical controls. Where the therapy to be tested is a local or systemic adjunct to SRP, a double-blind placebo-controlled randomized design in which one group receives SRP plus active agent, and the other receives SRP plus matched placebo, is tested (Jeffcoat 2002).

Another example of an ethics issues surrounding the use of placebo controls in periodontal

clinical trials is provided by Pihlstrom and Barnett (2010). If an established treatment has been shown to be efficacious or even is widely accepted for the treatment of a disease or condition, it would be considered unethical to withhold that treatment from trial participants. When a proven or established treatment is available, and a new treatment is to be evaluated, a trial would be designed to evaluate the superiority, equivalence, or non-inferiority of the new treatment relative to the established or 'proven effective' treatment. The periodontal treatment and pregnancy outcome trial protocols used as examples were approved by institutional review boards (IRBs) because periodontal treatment was not the standard of care for pregnant women when the trials were initiated. Thus, because the OPT trial (Michalowicz et al. 2006) and Smile Study (Newnham et al. 2009) showed that periodontal therapy was effective in reducing signs of periodontal disease in pregnant women, an IRB might now consider it unethical to conduct a trial that withholds periodontal therapy until after delivery in the control arm. In contrast, if the duration of withholding periodontal therapy is relatively short (6–9 months) and all participants are closely monitored for loss of periodontal support (clinical attachment level) and periodontal therapy is administered if there is progressive clinical attachment loss, use of a no-treatment control group in this situation might be viewed as ethical (Pihlstrom and Barnett 2010).

Most of the mouthrinse studies encouraged patients to use their normal oral hygiene methods twice per day. In the toothpaste studies, the agent and placebo agents were the toothpaste; thus, most of these studies were more exacting in the instructions including time and frequency of brushing and returning unused study material to check for compliance. For that reason, reductions over time in the placebo groups of toothpaste studies were included since the instructions were similar to those provided in clinical practice and the compliance checks most likely exceed those in private practice. Thus, any reduction in plaque or gingivitis in the placebo groups can be attributable to placebo effects and oral hygiene instructions and should mimic efficacy of oral hygiene instructions provided in clinical practice. Interestingly, the addi-

tional percent reductions due to the active agents in mouthrinses were similar in size to the reductions due to oral hygiene instructions aimed at improved mechanical plaque control (Gunsolley 2010).

4.2.3.4 Ethical Assessment of Trials

Weingarten et al. (2004) proposed that systematic reviews of experimental clinical research on humans should also include information on the ethical standards of the trials. The main reason for including ethics in the checklist of systematic reviews is to increase awareness in the scientific community about the need for high ethical standards in research on humans. The proposal would also encourage reviewers to identify those occasional studies that were so unethical that there may be doubts about the morality of using the results (Table 4.16).

Ethical Assessment: Examples

- *The study was approved by the Ethics Committee of the Eastman Dental Hospital (Reference 01/E012), and informed consent was obtained from all patients (Needleman et al. 2007).*
- *The present study was designed as a two-arm parallel-grouped, double-masked, placebo-controlled randomized clinical trial and was conducted in full accordance with the Helsinki Declaration of 1975, as revised in 2000. It has also been previously reviewed and approved by the Human Research Committee/Institute for Community Health Studies - Federal University of Rio de Janeiro (CEP/IESC - UFRJ, protocol #45/2007). To be enrolled, patients were individually informed about the nature of the proposed treatment, its risks and benefits, and signed informed consent forms. Additionally, every patient was granted the right to drop out the study, at any time, without jeopardizing the continuity of his/her periodontal treatment (Varela et al. 2011).*

4.2.4 Termination of a Clinical Trial

An RCT should not be continued longer than necessary since subjects enrolled in the group receiving the less advantageous therapy will be

Table 4.16 Guide for ethical assessment of trials in systematic reviews (Weingarten et al. 2004) (Reprinted with permission from BMJ Publishing Group Ltd.)

Goal related considerations
• Is there a clear declaration on financial support in all trials? • Is there a statement that relates to potential conflicts of interest in all trials? • Justification—Could the results have been obtained by laboratory or animal experiments? Were any of the trials superfluous? Was the size of the study sufficient to achieve adequate statistical power? • Publication bias—How many of the identified trials remained unpublished? Is bias detectable by funnel plot analysis?
Duty related considerations
• Were the comparators appropriate? If a placebo was used, was it justified?
Rights related considerations
• Safety—Was the risk for participants appropriate to the importance of the research? Was appropriate follow up care assured? • Was informed consent obtained? • When participants had reduced competence, were appropriate measures taken to protect their best interests? • Were adequate steps taken to prevent unauthorised access to personal and clinical data?
Global considerations
• Was the study approved by a research ethics committee?

harm. Further, other patients with the same disease who are not enrolled in the RCT will be deprived of the more advantageous therapy. Conversely, premature termination of an RCT may mean that data adequate to establish the safety and efficacy of test therapies may never be developed. Thus, rules for stopping an RCT should be developed prior to initiation of the trial. An attempt should be made to anticipate all possible contingencies (Dennison 1997).

There are four major **reasons for stopping a randomized clinical trial (RCT) early** (Bassler et al. 2008):

1. First, the trial may show serious adverse effects and may be stopped for unacceptable safety.
2. Second, investigators may stop the trial if interim differences in outcomes between the intervention and control groups are so unimpressive that any prospect of a positive result with the planned sample size is extremely unlikely—so-called ‘futility’. In this case, the argument goes, continuing the trial may not be justifiable in terms of time, money, and effort.
3. Third, new external information may arise during the conduct of the study that either convincingly answers the primary study question or that raises serious safety issues.

4. Finally, investigators may terminate an RCT for apparent benefit, the issue we will focus on in this article. Consequently, the investigators may conclude that one treatment is superior to the other and stop the trial before reaching the target sample size. Unfortunately, though generating low P values, these apparent effects may be due to chance. Random fluctuations will occur early in RCTs, and premature termination on the basis of these results risks attribution of these random fluctuations to true underlying treatment effects (Bassler et al. 2007).

Statistical modelling suggests that RCTs stopped early for benefit (**truncated RCTs**) will systematically overestimate treatment effects, and empirical data demonstrate that truncated RCTs often show implausibly large treatment effects (Montori et al. 2005; Bassler et al. 2010). Large differences in treatment effect size between truncated and nontruncated RCTs (ratio of relative risks <0.75) occurred with truncated RCTs having fewer than 500 events (Bassler et al. 2010).

Most systematic reviews and meta-analyses including truncated RCTs (tRCTs) fail to consider the possible overestimates of effect that

may result from early stopping for benefit (Bassler et al. 2007). Of 96 systematic reviews searched the Cochrane Library and MEDLINE that included at least one tRCT, 44 (46%) included 01 tRCT, 68 (71%) did not mention truncation at all, and 2 (2%) documented early stopping for benefit as a criterion for methodological quality. Of 47 meta-analyses in which authors reported, or were Bassler et al. (2007) could calculate the contribution of the tRCTs to the pooled result, the tRCTs contributed more than 40% of the weight in 16/47 (34%) (Bassler et al. 2007).

4.3 Consolidated Standards of Reporting Trials (Consort) Guidelines

As practitioners strive to improve the delivery of patient care, it is increasingly important that they find it easier to identify studies with the highest level of evidence. Once found, there must be some assurance that the studies were carried out satisfactorily and the methodology was sound. To assist in meeting this challenge, the **Consolidated Standards of Reporting Trials (CONSORT) guidelines** were developed by a team of dedicated journal editors, epidemiologists, and statisticians (Moher et al. 2001b). They determined the standards for authors reporting the findings of controlled clinical trials. CONSORT comprises a checklist and a flow diagram (Table 4.17 and Fig. 4.6) to help improve the quality of reports of randomized controlled trials (RCTs). It offers a standard way for researchers to report their findings (Turpin et al. 2005).

The checklist includes items, based on evidence, that should be addressed in the report. A lack of clarification of the meaning and rationale for each checklist item in the original CONSORT Statement has been remedied with the development of the CONSORT explanation and elaboration document (Altman et al. 2001; Schulz et al. 2010), which can also be found on the CONSORT Internet site (<http://www.consort-statement.org>). **The flow diagram** gives readers a clear picture of the progress of all participants in the trial, from the time

they are randomized until the end of their involvement. The intent is to make the experimental process transparent, flawed or not, so that users of the data can more appropriately evaluate the validity for their purposes. The key elements start when the researcher submits an article for review and completes the standardized forms. The completed checklist should accompany the manuscript through the review process and identify the page on which each item is addressed. The completed flow diagram should appear as a figure in the manuscript (Turpin et al. 2005). The CONSORT group also discussed the benefits of developing an explanatory document to enhance the use and dissemination of CONSORT. The document is patterned on reporting of statistical aspects of clinical research (Bailar and Mosteller 1988), which was developed to help facilitate the recommendations of the ICMJE's *Uniform Requirements for Manuscripts Submitted to Biomedical Journals* (Moher et al. 2001b).

4.3.1 Extended CONSORT Statement

The CONSORT Statement has been extended to cover:

- Different designs, such as **non-inferiority and equivalence trials** (Piaggio et al. 2006), **cluster randomized trials** (Campbell et al. 2004), and **pragmatic trials** (a pragmatic trial can be broadly defined as a randomized controlled trial whose purpose is to inform decisions about practice) (Zwarenstein et al. 2008)
- Types of interventions, such as **herbal therapies** (Gagnier et al. 2006), **non-pharmacologic treatments** (e.g., surgery, technical interventions, devices, rehabilitation, psychotherapy, and behavioural intervention) (Boutron et al. 2008), and acupuncture trials (The Standards for Reporting Interventions in Clinical Trials of Acupuncture: **STRICTA**) (MacPherson et al. 2010a, b; Hugh et al. 2010)
- Data, such as the **reporting of harms** (Ioannidis et al. 2004)
- Journals and conference **abstracts** (Hopewell et al. 2008b)

Table 4.17 CONSORT 2010 checklist of information to include when reporting a randomised trial (Schulz et al. 2010) (Reprinted with permission from BMJ Publishing Group Ltd. Available at <http://www.consort-statement.org>)

Section/topic	Item no	Checklist item	Reported on page no
Title and abstract			
	1a	Identification as a randomised trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
• Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
• Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
• Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	

Table 4.17 (continued)

Section/topic	Item no	Checklist item	Reported on page no
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	
<i>Discussion</i>			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
<i>Other information</i>			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

^aWe strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomized trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org

The CONSORT Statement focused on reporting parallel group-randomized trials in which individual participants are randomly assigned to study groups. However, in some situations, it is preferable to randomly assign groups of individuals (such as families or medical practices) rather than individuals. Reasons include the threat of contamination of some interventions (such as dietary interventions) if individual randomization is used. Also, in certain settings, randomization by group may be the only feasible method of conducting a trial. Trials with this design are variously known as field trials, community-based trials, place-based trials, or **cluster-randomized trials**. To accommodate the reporting of the spe-

cial features of the cluster-randomized trial, the CONSORT Statement was extended to include the following information:

- The rationale for adopting a cluster design
- How the effects of clustering were incorporated into the sample size calculations
- How the effects of clustering were incorporated into the analysis
- The flow of both clusters and individuals through the trial, from assignment to analysis (Campbell et al. 2004).

The Standards for Reporting Interventions in Clinical Trials of Acupuncture (**STRICTA**) were published in five journals in 2001 and 2002. These guidelines, in the form of a checklist and

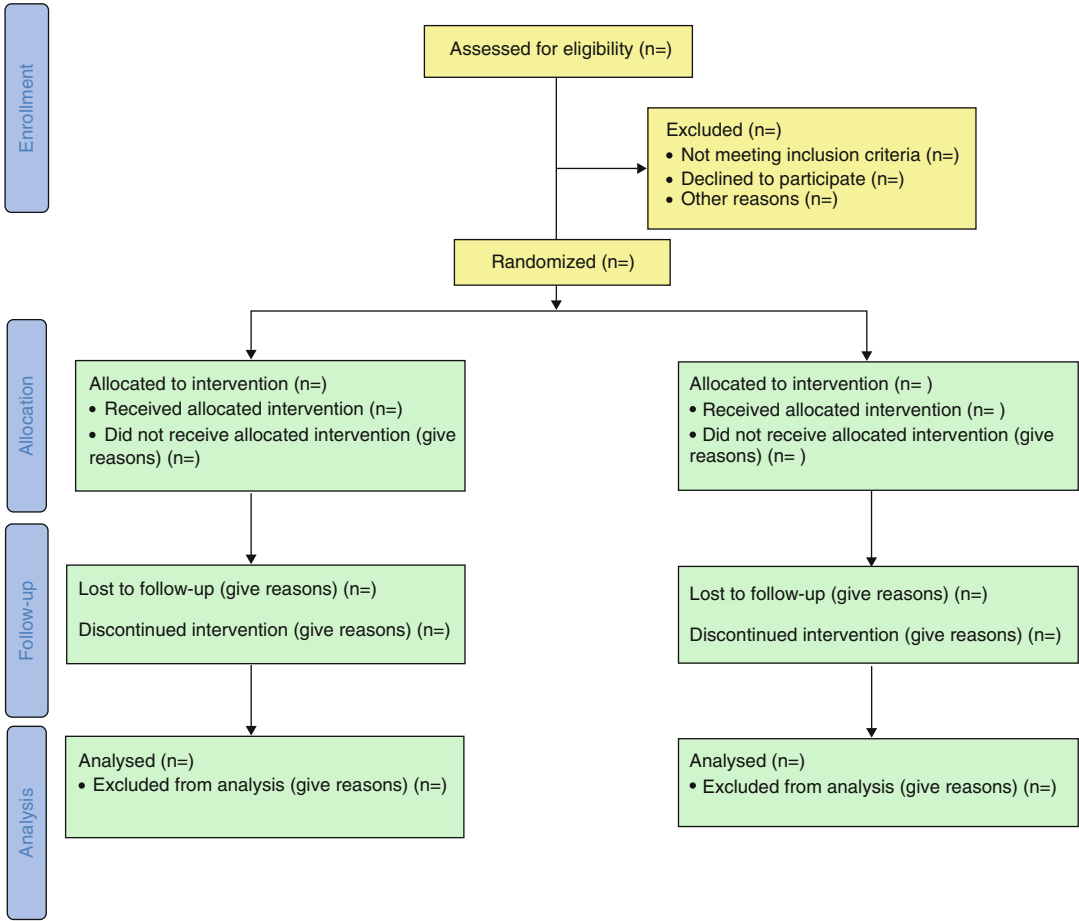


Fig. 4.6 CONSORT 2010 flow diagram of the progress through the phases of a parallel randomized trial of two groups (i.e., enrolment, intervention allocation, follow-

up, and data analysis) (Schulz et al. 2010. Reprinted with permission from BMJ Publishing Group Ltd.)

explanations for use by authors and journal editors, were designed to improve reporting of acupuncture trials, particularly the interventions, thereby facilitating their interpretation and replication. Subsequent reviews of the application and impact of STRICTA have highlighted the value of STRICTA as well as scope for improvements and revision. To manage the revision process, a collaboration between the STRICTA Group, the CONSORT Group, and the Chinese Cochrane Centre was developed in 2008. An expert panel with 47 participants was convened that provided electronic feedback on a revised draft of the checklist. At a subsequent face-to-face meeting in Freiburg, a group of 21 participants further revised the STRICTA checklist and planned dissemina-

tion. The new STRICTA checklist, which is an official extension of CONSORT, includes 6 items and 17 sub-items. These set out reporting guidelines for the acupuncture rationale, the details of needling, the treatment regimen, other components of treatment, the practitioner background, and the control or comparator interventions. In addition, and as part of this revision process, the explanations for each item have been elaborated, and examples of good reporting for each item are provided. In addition, the word ‘controlled’ in STRICTA is replaced by ‘clinical’, to indicate that STRICTA is applicable to a broad range of clinical evaluation designs, including uncontrolled outcome studies and case reports (MacPherson et al. 2010a, b).

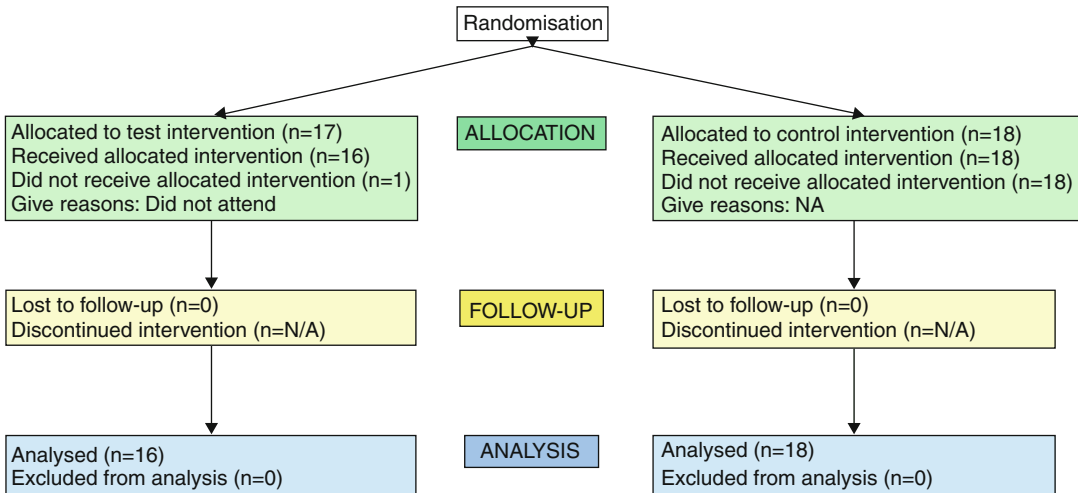


Fig. 4.7 Patient flow through study (Needleman et al. 2007). Reprinted with permission from John Wiley & Sons, Inc.)

Flow Diagrams: Examples

- Thirty-five subjects were recruited, although one withdrew before allocation of treatment group and was therefore not included in the study (Fig. 4.7). Consequently, 34 subjects were included in the analysis (18 control, 16 test), with an age range of 32–58 years at baseline. Four subjects did not complete the study (two in each group). Three subjects did not return due to personal problems and one due to loss of interest in the study. An intention-to-treat analysis used a last observation carried forward (Needleman et al. 2007).

4.3.2 Quality of Reporting Randomized Controlled Trials (RCTs) in the Medical Literature

Ideally, research publications need to convey to their readers relevant information about the design, conduct, analysis, and applicability of the results of trials. Lack of transparency in published reports is a barrier to accurate appraisal of the validity of trial results and may result in the use of ineffective or harmful treatments in clinical practice. The issue of the number of guidelines and extensions that have been published is also worthy of consideration. Although this diversity may be regarded as a strength as it produces guidelines tailored to most situations, it is also a potential

weakness as authors and editors of medical journals may be uncertain about which are the most appropriate guidelines to use. Even worse, they may feel obliged to try to adhere to more than one guideline or even abandon guidelines altogether. To overcome these issues, the **Enhancing the Quality and Transparency of Health Research (EQUATOR) Network** (<http://www.equator-network.org>), an international initiative that seeks ‘to improve the quality of scientific publications by promoting transparent and accurate reporting of health research’, was created in 2006. This network has three major objectives: to map the current status of all activities aimed at preparing and disseminating guidelines on reporting health research studies, to identify key individuals working in the area, and to establish relationships with potential key stakeholders (Boutron and Ravaud 2009).

Altman (2005) noted that in 2003, CONSORT was mentioned in the instructions of only 22% medical journals with high-impact factor, more often in general and internal medicine journals (53%) than in specialty journals (18%). However, 9/36 journals referred only to the obsolete 1996 statement, whereas the other 27 journals (16% of the sample) referred to the latest version gave the web address (<http://www.consort-statement.org/>) or both. No journal cited alternative reporting recommendations for randomized controlled trials. In 2008, Hopewell et al. (2008a) assessed again the extent to which leading medical journals

that publish reports of randomized trials incorporate the CONSORT recommendations into their journal and editorial processes. Thirty-eight percent of journals mentioned the CONSORT Statement in their online 'Instructions to Authors'; of these 37% stated this was a requirement, 63% were less clear in their recommendations. Very few journals mentioned the CONSORT extension papers. Journals that referred to CONSORT were more likely to refer to ICMJE guidelines (RR 2.16; 95% CI 1.51–3.08) and clinical trial registration (RR 3.67; 95% CI 2.36–5.71) than those journals which did not. 88% (50/57) of journals recommended authors comply with the CONSORT Statement; 62% said they would require this. 41% reported incorporating CONSORT into their peer-review process and 47% into their editorial process. 81% reported including CONSORT in their 'Instructions to Authors' although there was some inconsistency when cross checking information on the journal's website. 69% of journals recommended authors comply with the CONSORT extension for cluster trials, 60% for harms, and 42% for non-inferiority and equivalence trials.

Preliminary indications are that the use of CONSORT does indeed help to improve the quality of reports of RCTs (Moher et al. 2001b; Egger et al. 2001; Devereaux et al. 2002). In an evaluation, of 71 published RCTs, in three journals in 1994, allocation concealment was not clearly reported in 61% ($n=43$) of the RCTs (Moher et al. 2001a). Four years later, after these three journals required authors reporting an RCT to use CONSORT, the proportion of papers in which allocation concealment was not clearly reported had dropped to 39% (30 of 77, mean difference = -22% ; 95% confidence interval of the difference: -38% , -6%) (Egger et al. 2001). In a study using 29 medical journals, including those using and not using the CONSORT Statement, Devereaux et al. (2002) found that 60% of the RCT articles reported random sequence generation; 46% allocation concealment; 68% number of patients lost to follow-up; 64% method of analysis (intention to treat or other); 72% participant blinding status; 34% health-care provider blinding status; 37% data collector blinding sta-

tus; 48% judicial assessor blinding status; 4% data analyst blinding status; 93% baseline characteristics for each group; and 30% co-interventions for each group during the study. In addition, Devereaux's study showed that while promotion of the CONSORT checklist is indeed associated with improved reporting, there remains suboptimal reporting even among CONSORT-promoting journals (an average of 6.4 of 11 key methodological factors in an analysis that adjusted for the impact factor of the journal).

4.3.3 Quality of Reporting Randomized Controlled Trials (RCTs) in Dental Journals

Randomized controlled trials (RCTs) are regarded as the best study designs to test the efficacy of medical and dental intervention. Many reports, however, have shown that at the moment, the quality of dental RCT reports is still poor, and further efforts to improve it are necessary. It has been suggested that trials that are not well designed provide biased estimates of the treatment effects and that a journal's impact factor is not related to the quality of RCTs published (Cioffi and Farella 2011).

In a study assessing the quality of RCTs published by the end of 1999 concerned with the effectiveness of oral implants, it was revealed that randomization and concealment allocation procedures were not described in 70% articles, reasons for withdrawals were not given in 23% reports, and no attempt at blinding was reported in 72% of studies. The authors concluded that the quality of RCTs of oral implants is generally poor and needs to be improved (Esposito et al. 2001). Also, in a study by Dumbrigue et al. (2006), the quality of **implant dentistry** RCTs was assessed over a 10-year period (1991–2000), based on the reporting of control of potential sources of bias in the design methodology. Method of randomization was explicit in 51% of the RCTs, but only 12% incorporated blinding in the assessment of outcome. 98% accounted for all subjects at the end of the study. Looking at overall quality scores, only 2% of RCTs adequately reported on control of bias

in the three areas examined, 56% were deficient in one area, and 42% were deficient in two areas.

Recently, Marshman and Farid (2010) assessed the quality of the reporting of RCTs in **dental public health** journals (*Community Dental Health*, *Community Dentistry & Oral Epidemiology* and the *Journal of Public Health Dentistry*) over the period 1993–2008. Each of the resulting papers was reviewed and scored, according to 56 criteria, based on the CONSORT Statement. The search identified 48 papers. The average number of criteria present per article was 27.0 (SD=6.9), with variation between journals as follows: *Community Dental Health* (27.7), *Community Dentistry & Oral Epidemiology* (27.4), and *Journal of Public Health Dentistry* (23.2). The average number of criteria present per article increased over the time period used. The quality of the reporting could be improved if the CONSORT Statement was followed more closely.

The quality of RCTs in **prosthodontics**, excluding endosseous implant-based prosthetics, was evaluated by Jokstad et al. (2002). Randomization and procedures for concealment allocation were not described in 70% of the articles. The methods used to generate the random allocation sequence were not mentioned in 82%. The methods used to implement the random allocation sequence, clarifying whether it was concealed until all interventions were assigned, were not mentioned in 94%. Reporting who generated allocation sequence, who enrolled patients, and who assigned participants to groups was not reported in 7%. Reasons for withdrawals were not given in 23% of the reports. No attempt at blinding was reported in 72%. Statistical analysis was not described in 6% of the papers, while these analyses were assessed as appropriate for 75%, unclear in 12%, and inappropriate in 7%.

Montenegro et al. (2002) assessed the quality of RCTs in **periodontology** using these evidence-based components. The results showed that although 91% of trials were described as randomized, adequate methods for randomization and allocation concealment were found in 17% and 7% of studies, respectively. Blinding was adequate for the caregiver in 17% and for the examiner in 55% of studies. A clear accounting of all

participants was present in 56% of reports. This rigorous systematic review revealed that the quality of RCTs in periodontology, judged by their publications, frequently does not meet recommended standards.

So far, the quality of reporting **orthodontic** clinical trials published between 1989 and 1998 was insufficient to allow readers to assess the validity of the trials. One hundred and fifty-five trial reports were identified of which 2.6% were adequately concealed, 54.8% were described as being randomized, 6.5% as double-blind, and 28.4% gave a description of withdrawals and dropouts from the trial. The type of randomization was considered appropriate in 50.3% reports, and in 36.8% reports, the level of blinding was considered appropriate. When assessed for the risk of bias in the reported trials, one trial (0.6%) had a low risk of bias, 17 (11%) a moderate risk, and 137 (88.4%) a high risk. In orthodontics, it is often very difficult to carry out triple- or even double-blind trials because orthodontic appliances and materials often differ in appearance so that participants and/or clinicians are aware of which intervention any participant is receiving. For example, in bonding trials, is it possible for patients or clinicians to detect which bracket is bonded with which material—are they slightly different colours; were they cured in different ways, for example, one light cured and the other chemical? If so, who should assess whether a bracket is off or a band loose? In this situation, it may not be possible to achieve any level of blinding, so open trials may be appropriate. For other studies, for example, comparisons of functional appliances, it is probably not possible to blind patients or the clinicians who are treating them, but records can be assessed independently with the assessor blind to the intervention used. So, for these trials, an appropriate level of blinding would be where records are anonymous and examined by an independent assessor away from the participants. However, it is possible to adopt double- or triple-blinding strategies in studies that assess different mouthwashes, toothpastes, or analgesics that can be prepared and packaged to be identical (Harrison 2003).

In a report concerning RCTs in **paediatric dentistry** (Al-Namankany et al. 2009), the quality

Table 4.18 The modified 34-item Consolidated Standards of Reporting Trials checklist used by Al-Namankany et al. (2009) (Al-Namankany et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Article section and topic	
Title and abstract	1. Were the words ‘random allocation’, ‘randomized’, or ‘randomly assigned’ mentioned in the abstract?
Background	2. Were the nature, scope, and severity of the problem described?
Participants	3. Were the eligibility (and exclusion) criteria of the trial participants described: age, gender, clinical diagnosis, and comorbid conditions? 4. Were the settings and locations of the data collection reported?
Interventions	5. Were precise details of the interventions intended for each group given? Was it described how and when they were actually administered?
Objectives	6. Were the specific objectives and/or hypotheses mentioned?
Outcomes	7. Were there clearly defined primary and secondary outcome measures (provenance and properties of the scales), and did they distinguish between first- and second-degree outcomes? 8. Were any methods used to enhance the quality of measurements (e.g. multiple observations, training of assessors)?
Sample size	9. Is there a description of how sample size was calculated, identification of the primary outcome on which the calculation was based, and the resulting target sample size per comparison group? 10. Did they provide any explanation on interim analyses or stopping rules?
Randomization sequence generation	11. Was the method used to generate the random allocation sequence reported? 12. If applicable, did the authors provide details of any restriction (e.g. blocking, stratification)?
Randomization allocation concealment	13. Was the method used to implement the random allocation sequence (e.g. numbered containers or central telephone) reported, clarifying whether the sequence was concealed until interventions were assigned?
Randomization implementation	14. Did the authors explain who generated the allocation sequence? 15. Did the authors explain who enrolled the participants? 16. Did the authors explain who assigned the participants to their groups?
Blinding (masking)	17. Did the authors report whether participants were blinded to group assignment? 18. Did the authors report whether those administering the interventions were blinded to group assignment? 19. Did the authors report whether the assessors were blinded to group assignment? 20. If blinded, did they report how the success of blinding was evaluated?
Statistical methods	21. Did the authors specify which statistical procedure was used for primary outcomes? 22. When applicable, were the methods of subgroup analysis or adjusted analyses given, and/or did the authors clarify the choice of variables that were adjusted for and specify whether the analysis was planned or suggested by the data?
Participant flow	23. Did the authors report the flow of each participant through each stage (or if there is no flow chart included in the study report, is it possible to fill into the CONSORT flow diagram the number of participants randomly as signed to each group, receiving intended treatment, completing the study protocol, and analysed for primary outcome)? 24. Did the authors describe study protocol violations together with reasons?
Recruitment	25. Are the dates defining the periods of recruitment and follow-up given?
Baseline data	26. Are the baseline demographic and clinical characteristics of each group (mean, SD, or categories) reported?
Numbers analysed	27. Is the number of participants (denominator) in each group included in each analysis (e.g. in binary outcomes, the results should be stated in absolute numbers, not in proportions)? 28. Was it an ‘intention-to-treat’ analysis (the participants were analysed in the same groups to which they were randomized)?
Outcomes and estimation	29. Is there a summary of results and the estimated effect size and precision (e.g. 95% confidence interval) for each group reported for each primary and secondary outcome?

Table 4.18 (continued)

Article section and topic	
Ancillary analyses	30. When undertaken, did the authors indicate which ancillary analyses (subgroup or adjusted analysis) were prespecified in the protocol?
Adverse events	31. Are the estimates of the frequency of all important adverse or side effects in each intervention group reported?
Discussion – interpretation	32. Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes.
Generalizability	33. Generalizability (external validity) of the trial findings.
Overall evidence	34. General interpretation of the results in the context of current evidence.

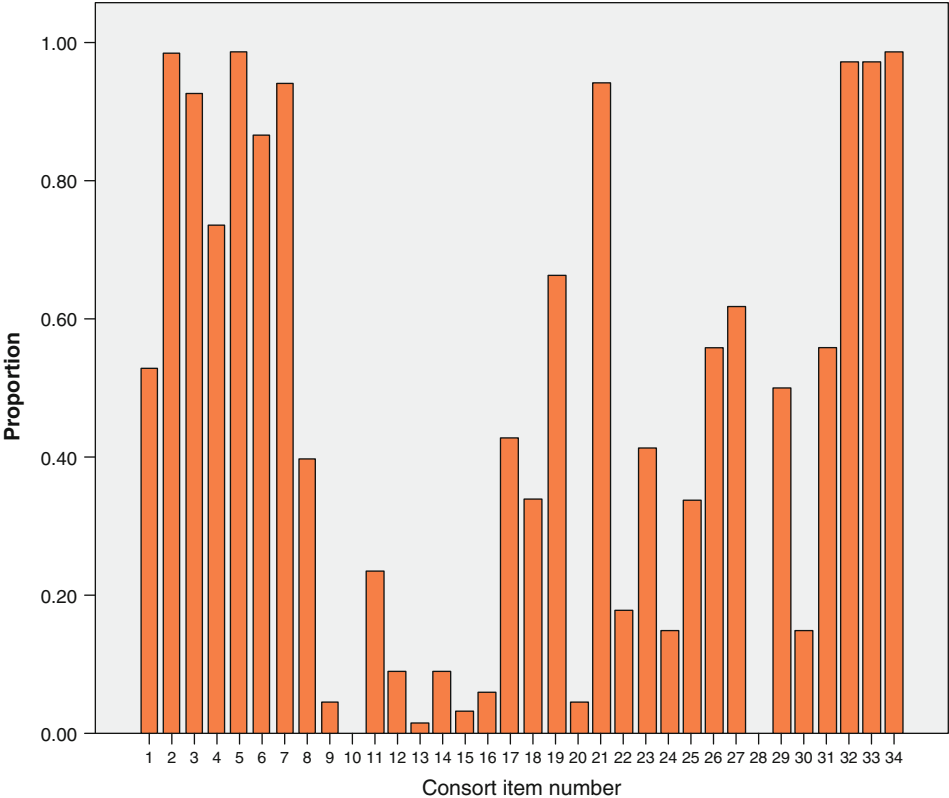


Fig. 4.8 Compliance of articles in paediatric dentistry against the Consolidated Standards of Reporting Trials (CONSORT) checklist (Al-Namankany et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

of trials published between 1985 and 1997, and between 1998 and 2006 were compared to assess the influence of CONSORT introduction on quality of published RCTs. The compliance assessment was undertaken using an operational version of the CONSORT checklist (Table 4.18), whereby the 22 items of the CONSORT Statement checklist were converted into 34 questions. Good levels of compliance with CONSORT were found in the

reporting of the introductions, which included the title, abstract, and background of the studies (96–98%), and also in the discussion, interpretation, generalizability, and overall evidence sections (97%, 98%, and 96%, respectively) (Fig. 4.8). Reporting of randomization methods was poor, ranging from 5% to 9% compliance. Only 8 of 173 (4%) of published articles provided a description of how the sample size was calculated. An

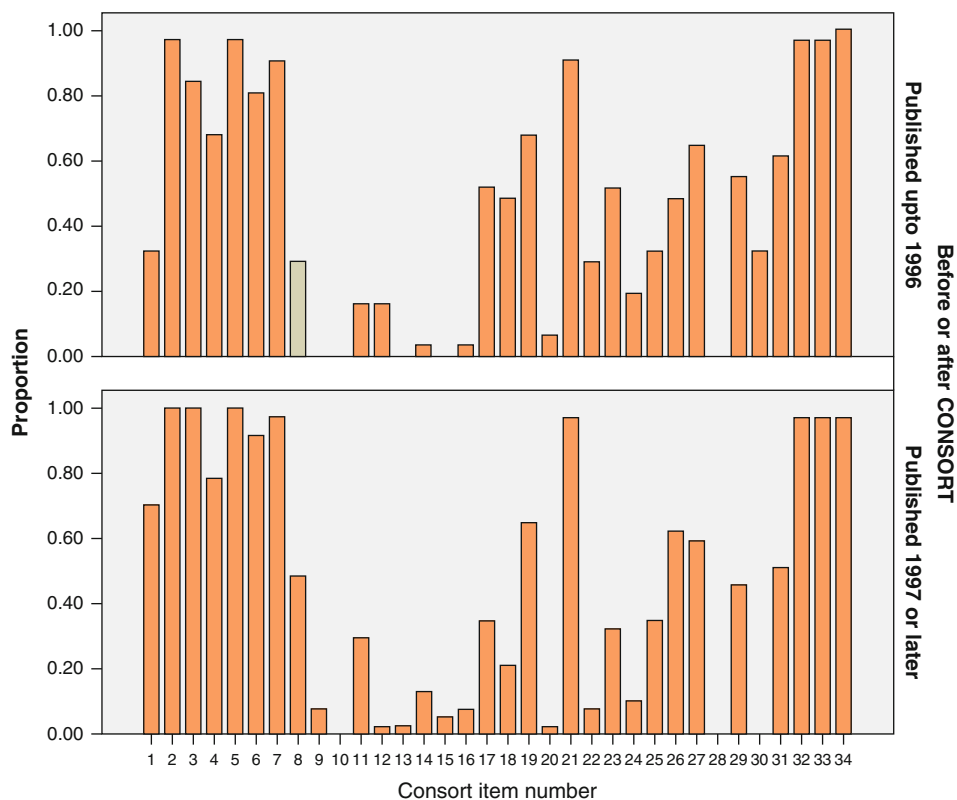


Fig. 4.9 Compliance with Consolidated Standards of Reporting Trials (CONSORT) in paediatric dentistry before and after 1997 (Al-Namankany et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

‘intention-to-treat’ analysis was only reported in 2 of 173 (1%) published articles. The quality of reporting showed better results after 1997 in the titles/abstracts and discussion sections. Compliance with item 18, which asked if authors report whether those administering the interventions were masked (blinded) to group assignment, worsened. There was little change in other checklist items pre- and post-CONSORT. A comparison of compliance with the CONSORT checklist until 1996 and after 1997 is illustrated in **Fig. 4.9**.

The quality of published RCTs has been examined also in relation to journal impact factor in different areas of dental research by Sjögren and Halling (2002). The randomization process was judged to be adequate in only 27% of cases, and blinding was reported in just 28% of papers. The accounting for withdrawals/dropouts was present in 35% of cases, and allocation concealment was reported in 13% of the trials. No cor-

relation between the journal impact factor and the quality of RCTs was found. This result has also been confirmed by Barbui et al. (2006) who showed that the impact factor of journals is not a valid measure of randomized controlled trial quality.

When the quality of reporting of RCTs published in dental specialty journals possessing the highest impact factor (2008 data) was investigated, 95 RCTs were identified with generally suboptimal scores on quality reporting on key CONSORT areas. Significant differences were found among journals with the *Journal of Clinical Periodontology* achieving the highest score, followed by the *American Journal of Orthodontics* and *Dentofacial Orthopedics*. There was a positive association between quality score and number of authors, involvement of statistician/epidemiologist, and multicentre trials (Pandis et al. 2010).

Table 4.19 Examples of deficiencies in clinical trial reports cited by peer reviewers (Pihlstrom and Barnett 2010) (Reprinted with permission from Sage Publications)

• Specific inclusion criteria were not listed or were ambiguous.
• Participants were included who did not meet inclusion criteria.
• The method of group assignment was not presented.
• Description of the study design was incomplete and/or procedures were not described clearly.
• The study objective was not stated.
• No basis for the sample size was presented.
• There was insufficient description of the clinical assessment method.
• A statistical analysis of subgroups but not of the entire study population was presented.
• Calibration and reproducibility of multiple examiners were not described.
• The paper contained multiple, inconsistent descriptions of study details, e.g., the claimed examination intervals were different in the abstract, materials and methods, and results sections.
• Lack of controls or inadequate controls.
• Conclusions not warranted by the data, e.g., statement of treatment ‘equivalence’ when study and statistical analysis were not designed for this purpose.
• Incomplete statistical section; primary and secondary variables not identified.

As a condition for publication of clinical trials, Pihlstrom (2009) recommended that high-impact oral health journals do the following:

- Require adherence to the CONSORT Statement recommendations.
- Unless specified otherwise in a public trial registry before enrolling subjects, require an ITT analysis or some other acceptable means of data imputation for missing data in RCTs.
- Require registration of phase III RCTs on a public website before randomization.

Adoption of these conditions for publication will provide clear and transparent guidance to authors and reviewers and will give policymakers and clinicians the unambiguous information they need to make sound decisions based on evidence from RCTs (Pihlstrom 2009).

A sampling of prominent deficiencies cited in the peer review of submitted papers is given in **Table 4.19**. While some of these deficiencies might be dealt with in revisions, others are severe enough to preclude acceptance for publication since they cannot be rectified after the study has been concluded. For the most part, a review of the CONSORT criteria at the time the study was designed and/or the manuscript written might have solved these problems (Pihlstrom and Barnett 2010).

Improved reporting quality would be a major advance in biomedical research, offering the potential to accelerate advances in health care. Compared with the actual development and conduct of research studies, improving reporting quality is relatively straightforward and has modest requirements for resources. In addition, change could be brought about rapidly. A concerted and shared approach by stakeholders would have a profound impact on the quality of publications, maximizing benefits to consumers and other stakeholders reading research reports. This should be a shared obligation and responsibility (Needleman et al. 2008).

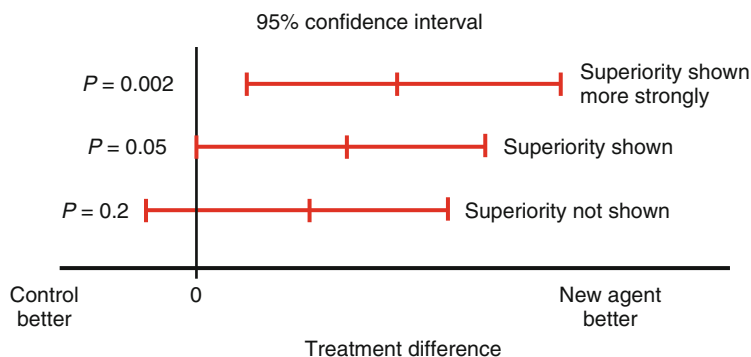
4.4 Special Types of RCTs: Comparative RCTs

An important issue in evaluating the quality of RCTs is the distinction between superiority and equivalence (Tu et al. 2006).

4.4.1 Superiority Trial

A superiority trial is designed to detect a difference between treatments. The first step of the

Fig. 4.10 Relationship between significance tests and confidence intervals (Committee for Proprietary Medicinal Products 2001. Reprinted with permission from John Wiley & Sons, Inc.)



analysis is usually a test of statistical significance to evaluate whether the results of the trial are consistent with the assumption of there being no difference in the clinical effect of the two treatments. In a trial of good quality, the degree of statistical significance (*P* value) indicates the probability that the observed difference, or a larger one, could have arisen by chance, assuming that no difference really existed. Once it is accepted that the assumption of ‘no difference’ is untenable, it then becomes important to estimate the size of the difference in order to assess whether the effect is clinically relevant. This has two aspects. First, there is the best estimate of the size of the difference between treatments (**point estimate**). For normally distributed data, this is usually taken as the observed difference between the mean values on each. Second, there is the range of values of the true difference that are plausible in the light of the results of the trial (**confidence interval**). The method of constructing confidence intervals generally ensures that this is so, provided it corresponds to the choice of significance test. Thus, the following two statements are usually equivalent: (1) The two-sided 95% confidence interval for the difference between the means excludes zero; (2) The two means are statistically significantly different at the 5% level ($P < 0.05$) two-sided (**Fig. 4.10**) (Committee for Proprietary Medicinal Products 2001).

4.4.2 Equivalence Trial

An equivalence trial is designed to confirm the absence of a meaningful difference between treatments. In this case, it is more informative to

conduct the analysis by means of the calculation and examination of the confidence interval, although there are closely related methods using significance test procedures. A **margin of clinical equivalence** (Δ) is chosen by defining the largest difference that is clinically acceptable, so that a difference bigger than this would matter in practice. If the two treatments are to be declared equivalent, then the two-sided 95% confidence interval—which defines the range of plausible differences between the two treatments—should lie entirely within the interval $-\Delta$ to $+\Delta$, see **Fig. 4.11**. There are situations in which the equivalence margins may be chosen asymmetrically with respect to zero (Committee for Proprietary Medicinal Products 2001).

In the case of bioequivalence studies, a coverage probability of 90% for the confidence interval has become the accepted standard when evaluating whether the average values of the pharmacokinetic parameters of two formulations are sufficiently close (Committee for Proprietary Medicinal Products 2001).

Clinical equivalence trials, with two-sided 95% confidence intervals, may be carried out when conventional bioequivalence trials are impossible, for example, in the case of a generic inhaled or topically applied product (Committee for Proprietary Medicinal Products 2001).

4.4.3 Non-inferiority Trial

In phase III drug development, non-inferiority trials are more common than equivalence trials. In these, we wish to show that a new treatment is no less effective than an existing treatment—it

Fig. 4.11 Confidence interval approach to analysis of equivalence trial (Committee for Proprietary Medicinal Products 2001). Reprinted with permission from John Wiley & Sons, Inc.)

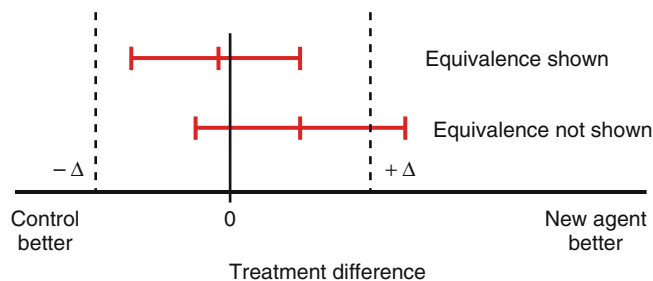
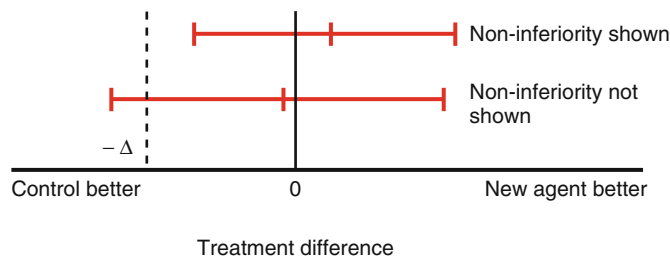


Fig. 4.12 Confidence interval approach to analysis of non-inferiority trial (Committee for Proprietary Medicinal Products 2001). Reprinted with permission from John Wiley & Sons, Inc.)



may be more effective or it may have a similar effect. Again, a confidence interval approach is the most straightforward way of performing the analysis, but now we are only interested in a possible difference in one direction. Hence, the two-sided 95% confidence interval should lie entirely to the right of the value $-\Delta$, see **Fig. 4.12**. Non-inferiority trials are sometimes mistakenly referred to, and designed as, equivalence trials. This distinction is important and can be a source of confusion (Committee for Proprietary Medicinal Products 2001).

Note also that by using the closely related significance testing procedures referred to **equivalence trial**, it is possible to calculate a P value associated with the null hypothesis of inferiority. This is a valuable further aid to assessing the strength of the evidence in favour of non-inferiority (Committee for Proprietary Medicinal Products 2001).

4.4.4 Interpreting a Non-inferiority Trial as a Superiority Trial

If the 95% confidence interval for the treatment effect not only lies entirely above $-\Delta$ but also above zero, then there is evidence of superiority in terms of statistical significance at the 5% level ($P < 0.05$). In a superiority trial, the full analysis

set, based on the ITT (intention-to-treat) principle, is the analysis set of choice, with appropriate support provided by the per-protocol analysis set. In a non-inferiority trial, the full analysis set and the per-protocol analysis set have equal importance, and their use should lead to similar conclusions for a robust interpretation. A switch of objective would require this difference of emphasis to be recognized.

Switching the objective of a trial from non-inferiority to superiority is feasible provided:

1. The trial has been properly designed and carried out in accordance with the strict requirements of a non-inferiority trial.
2. Actual P values for superiority are presented to allow independent assessment of the strength of the evidence.
3. Analysis according to the intention-to-treat principle is given greatest emphasis (Committee for Proprietary Medicinal Products 2001).

4.4.5 Interpreting a Superiority Trial as a Non-inferiority Trial

Switching the objective of a trial from superiority to non-inferiority may be feasible provided:

1. The non-inferiority margin with respect to the control treatment was predefined or can be

justified. (The latter is likely to prove difficult and to be limited to rare cases where there is a widely accepted value for Δ .)

2. Analysis according to the intention-to-treat principle and per-protocol analysis, showing confidence intervals and *P* values for the null hypothesis of inferiority, give similar findings.
3. The trial was properly designed and carried out in accordance with the strict requirements of a non-inferiority trial (see ICH E9 and E10).
4. The sensitivity of the trial is high enough to ensure that it is capable of detecting relevant differences if they exist.
5. There is direct or indirect evidence that the control treatment is showing its usual level of efficacy (Committee for Proprietary Medicinal Products 2001).

4.4.6 Interpreting a Superiority as Equivalence Trials

The research hypotheses and study designs are entirely different for these two dissimilar research strategies. For a superiority trial, the null hypothesis is that there is no difference between treatments. The trial is designed to reject the null hypothesis, thereby showing that the test group achieves a statistically better outcome. In contrast, equivalence trials test the null hypothesis that any difference that occurs between the treatments is greater than the pre-specified equivalence margin. If the results reject the null hypothesis, where the difference in treatment efficacy falls within the pre-specified margin, the new treatment is shown not to be inferior to or is 'as good as' the active control. A common error in superiority trials is to assume that the two treatments are equivalent if the results fail to reject the null hypothesis (Tu et al. 2006).

Also, the required sample size for an equivalence trial might be substantially greater than that for a superiority trial (Tu et al. 2005). Underpowered active-controlled trials might give rise to a false impression that the new treatment is as good as the established one, if there is no

distinction between superiority and equivalence in their research hypotheses (Tu et al. 2006).

In a superiority trial, a pre-specified margin of difference between the two treatments that is deemed large enough to be clinically significant has to be specified before the trial commences. In an equivalence trial, the equivalence margin that is deemed sufficiently small by clinicians not to differentiate the treatment effects as clinically significant must also be specified before the trial commences. For the latter, when the confidence interval of the difference in the efficacy between the treatments falls within the pre-specified margin, these two treatments are considered equivalent (Tu et al. 2006).

While most clinical studies are done to show comparative superiority, but many reports now claim equivalence between the investigated entities, Greene et al. (2000) assessed the justification and support for claims of clinical or therapeutic equivalence in medical journals. Only 51% of the reports were specifically aimed at studying equivalence; the others either tried to show superiority or did not state a research aim. The quantitative distinctions regarded as 'equivalent' ranged from 0% to 21% for direct increments and from 0% to 76% for proportionate differences. An equivalence boundary was set and confirmed with an appropriate statistical test in only 23% of reports. In 67% of reports, equivalence was declared after a failed test for comparative superiority, and in 10%, the claim of equivalence was not statistically evaluated. The sample size needed to confirm results had been calculated in advance for only 33% of reports. Sample size was 20 patients per group or fewer in 25% of reports. These methodologic flaws can lead to false claims, inconsistencies, and harm to patients.

Although superiority and equivalence trials share common characteristics in study design—for example, random allocation, concealment of randomization, blindness, intention to treat, etc.—there are additional considerations for study designs of equivalence trials when an active control is used (Greene et al. 2000; McAlister and Sackett 2001). This is important, because the design and interpretation of equivalence trials require a different methodology if the results and

Table 4.20 Criteria for deciding whether a trial is valid (McAlister and Sackett 2001) (Reprinted with permission from Elsevier)

Superiority trials	Active-control equivalence trials ^a
Was the assignment of patients to treatments randomized?	Was the assignment of patients to treatments randomized?
Was the randomization list concealed?	Was the randomization list concealed?
Were all patients who entered the trial accounted for at its conclusion?	Were all patients who entered the trial accounted for at its conclusion?
Was analysis by intention-to-treat?	Was analysis by intention-to-treat <i>and on-treatment</i> ?
Were clinicians and patients blinded to which treatment they received?	Were clinicians and patients blinded to which treatment they received?
Aside from the experimental treatment, were the two groups treated equally?	Aside from the experimental treatment, were the two groups treated equally?
Were the groups similar at baseline?	Were the groups similar at baseline?
Were the study endpoints clinically important (efficacy and safety)?	Were the study endpoints clinically important (efficacy and safety)?
	<i>Was the active control previously shown to be effective?</i>
	<i>Were study patients and outcome variables similar to those in the original trials establishing the efficacy of the active control?</i>
	<i>Were both regimens applied in optimal fashion?</i>
	<i>Was the appropriate null hypothesis tested?</i>
	<i>Was the equivalence margin specified before the study?</i>
Was the trial large enough?	Was the trial large enough?

^aCriteria that are unique to active-control equivalence trials are italicized

conclusions are to be valid (Table 4.20). Failure to show the superiority of one treatment over the other does not prove that these two treatments are equivalent (Tu et al. 2006).

4.4.7 Comparative RCTs in the dental literature

The dental literature is not short of discussions around the distinction between superiority and equivalence trials (Fleiss 1992; Kingman 1992; Listgarten 1992; Koch and Paquette 1997; Duke and Garrett 1998; Gunsolley et al. 1998; Burns and Elswick 2001). Nevertheless, dental research is still plagued by the misconception that if no difference in the treatment effect is found, then treatments are shown to be equivalent. A recent review assessing the quality of active-controlled trials in periodontal research (Tu et al. 2006) clearly showed that some researchers remain unaware that testing therapeutic equivalence needs a completely different approach from testing superiority, with regard to both null hypothe-

sis and sample size calculation (Tu et al. 2005). Failing to reject the null hypothesis in a superiority trial does not mean that the null hypothesis is true; it is plausible that a larger study would always reach a smaller P value, thereby rejecting the null hypothesis given that the difference in treatment effect remains the same. The confusion around the differences in objectives between superiority and equivalence trials might be responsible for the design of these trials and their inappropriate interpretation (Tu et al. 2006).

4.5 Special Types of RCTs: Cluster (Group)-Randomized Trials

GROUP-randomized trials (GRTs) are comparative studies designed to evaluate interventions that operate at a group level, manipulate the physical or social environment, or cannot be delivered to individuals. Examples include school-, work-site-, and community-based studies designed to improve the health of students, employees, and residents, respectively. Just as the randomized

clinical trial (RCT) is the gold standard in public health and medicine when allocation of individual participants is possible, the GRT is the gold standard when allocation of identifiable groups is necessary (Murray et al. 2004).

There are four characteristics that distinguish the GRT from the more familiar RCT:

1. First, the unit of assignment is an identifiable group; such groups are formed not at random but rather through some physical, social, geographic, or other connection among their members.
2. Second, different groups are assigned to each condition, creating a nested or hierarchical structure for the design and the data.
3. Third, the units of observation are members of those groups nested within both their condition and their group.
4. Fourth, usually only a limited number of groups are assigned to each condition.

The number of published GRTs continues to increase, as evidenced by the identification of 60 articles in 5 years (between 1998 and 2002) by Varnell et al. (2004), compared with the 21 articles found in 4 years in the review by Simpson et al. (1995) between 1990 and 1993.

Only a few of the cluster trials reported in the dental public health literature have used randomization in the assignment of groups to experimental conditions, and most of these have appeared in the last few years (Masouredis et al. 1997; Gordon et al. 2005; Bahrami et al. 2004; Kiang et al. 2005; Gansky et al. 2005; Finch et al. 2005; Blinkhorn et al. 2003). Most of these studies do not apply the appropriate principles for the design and analysis of cluster trials. The most important feature of cluster trials is the existence of intra-class correlation (ICC) caused by correlation among members of the same group. Ignoring the ICC can lead to false conclusions. Special consideration is needed in designing cluster trials and analysing data from them. Further, groups rarely are randomized to interventions in these trials (Kim et al. 2006). Also, the method of cluster randomization may result in a random effect and cluster selection bias, which has to be taken into account when analysing and interpreting the results (van der Putten et al. 2010).

References

- Abaid LN, Grimes DA, Schulz KF. Reducing publication bias of prospective clinical trials through trial registration. *Contraception*. 2007;76(5):339–41.
- Al-Namankany AA, Ashley P, Moles DR, et al. Assessment of the quality of reporting of randomized clinical trials in paediatric dentistry journals. *Int J Paediatr Dent*. 2009;19:318–24.
- Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. *BMJ*. 2005;330(7499):1056–7.
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al.; CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663–94.
- Altman DG. Statistics and ethics in medical research: III How large a sample? *Br Med J*. 1980;281(6251):1336–8.
- Antczak AA, Tang J, Chalmers TC. Quality assessment of randomized controlled trials in dental research. II. Results: periodontal research. *J Periodontal Res*. 1986;21:315–21.
- Bahrami M, Deery C, Clarkson JE, Pitts NB, Johnston M, Ricketts I, et al. Effectiveness of strategies to disseminate and implement clinical guidelines for the management of impacted and unerupted third molars in primary dental care, a cluster randomized controlled trial. *Br Dent J*. 2004;197:691–6.
- Bailar 3rd JC, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. *Ann Intern Med*. 1988;108:266–73.
- Bang H, Ni L, Davis C E. Assessment of blinding in clinical trials. *Controlled Clinical Trials*. 2004;25:143–156.
- Barbui C, Cipriani A, Malvini L, et al. Validity of the impact factor of journals as a measure of randomized controlled trial quality. *J Clin Psychiatry*. 2006; 67:37–40.
- Bassler D, Ferreira-Gonzalez I, Briel M, Cook DJ, Devereaux PJ, Heels-Ansdell D, Kirpalani H, Meade MO, Montori VM, Rozenberg A, Schünemann HJ, Guyatt GH. Systematic reviewers neglect bias that results from trials stopped early for benefit. *J Clin Epidemiol*. 2007;60:869–73.
- Bassler D, Montori VM, Briel M, Glasziou P, Guyatt G. Early stopping of randomized clinical trials for overt efficacy is problematic. *J Clin Epidemiol*. 2008;61:241–6.
- Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, Heels-Ansdell D, Walter SD, Guyatt GH; STOPIT-2 Study Group, Flynn DN, Elamin MB, Murad MH, Abu Elnour NO, Lampropoulos JF, Sood A, Mullan RJ, Erwin PJ, Bankhead CR, Perera R, Ruiz Culebro C, You JJ, Mulla SM, Kaur J, Nerenberg KA, Schünemann H, Cook DJ, Lutz K, Ribic CM, Vale N, Malaga G, Akl EA, Ferreira-Gonzalez I, Alonso-Coello P, Urrutia G, Kunz R, Bucher HC, Nordmann AJ, Raatz H, da Silva SA, Tuche F, Strahm B, Djulbegovic B, Adhikari NK, Mills EJ, Gwadry-Sridhar F, Kirpalani

- H, Soares HP, Karanickolas PJ, Burns KE, Vandvik PO, Coto-Yglesias F, Chrispim PP, Ramsay T. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303:1180–7.
- Beller EM, Gebiski V, Keech AC. Randomisation in clinical trials. *Med J Aust*. 2002;177:565–7.
- Blinkhorn AS, Gratrix D, Holloway PJ, Wainwright-Stringer YM. A cluster randomised, controlled trial of the value of dental health educators in general dental practice. *Br Dent J*. 2003;195:395–400.
- Boutron I, Ravaud P. Improving the reporting of clinical research. *Br J Surg*. 2009;96:449–50.
- Boutron I, Estellat C, Ravaud P. A review of blinding in randomized controlled trials found results inconsistent and questionable. *J Clin Epidemiol*. 2005;58:1220–6.
- Boutron I, Estellat C, Guittet L, Dechartres A, Sackett DL, Hrobjartsson A, et al. Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review. *PLoS Med*. 2006;3:e425.
- Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med*. 2007;4:e61.
- Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med*. 2008;148(4):295–309.
- Burns DR, Elswick RK. Equivalence testing with dental clinical trials. *J Dent Res*. 2001;80:1513–7.
- Cakir B, Gebiski VJ, Keech AC. Flow of participants in randomised studies. *Med J Aust*. 2003;178:347–9.
- Campbell MK, Elbourne DR, Altman DG, CONSORT Group. CONSORT statement: extension to cluster randomised trials. *BMJ*. 2004;328(7441):702–8.
- Carpenter Jr WT, Appelbaum PS, Levine RJ. The Declaration of Helsinki and clinical trials: a focus on placebo-controlled trials in schizophrenia. *Am J Psychiatry*. 2003;160:356–62.
- Chalmers TC. Letters to the editor. When should randomization begin? *Lancet*. 1968;1:858.
- Chatellier G, Zapletal E, Lemaitre D, Menard J, Degoulet P. The number needed to treat: a clinically useful nomogram in its proper context. *BMJ*. 1996;312(7028):426–9.
- Christensen E. Methodology of superiority vs. equivalence trials and non-inferiority trials. *J Hepatol*. 2007;46:947–54.
- Cioffi I, Farella M. Quality of randomised controlled trials in dentistry. *Int Dent J*. 2011;61:37–42.
- ClinicalTrials.gov. <http://www.clinicaltrials.gov>. Accessed 11 May 2011.
- Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Medical Stat Med. 1989;8:441–54.
- Committee for Proprietary Medicinal Products (CPMP). Points to consider on switching between superiority and non-inferiority. *Br J Clin Pharmacol*. 2001;52:223–8.
- Cooley WE, Castellion AW. International regulatory aspects of clinical periodontal research. *Ann Periodontol*. 1997;2:18–30.
- Current Controlled Clinical Trials. <http://www.controlled-trials.com/>. Accessed 11 May 2011.
- De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *N Engl J Med*. 2004a;351:1250–1.
- De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Lancet*. 2004b;364:911–2.
- De Angelis CD, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA*. 2004c;292:1363–4.
- Dennison DK. Components of a randomized clinical trial. *J Periodontol Res*. 1997;32:430–8.
- Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA*. 2001;285:2000–3.
- Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Control Clin Trials*. 2002;23:380–8.
- Doig GS, Simpson F. Randomization and allocation concealment: a practical guide for researchers. *J Crit Care*. 2005;20:187–91.
- Duke SP, Garrett S. Equivalence in periodontal trials: a description for the clinician. *J Periodontol*. 1998;69:650–4.
- Dumbrigue HB, Al-Bayat MI, Ng CC, et al. Assessment of bias in methodology for randomized controlled trials published on implant dentistry. *J Prosthodont*. 2006;15:257–63.
- Endacott R, Botti M. Clinical research 3: Sample selection. *Accid Emerg Nurs*. 2007;15(4):234–8.
- Egger M, Jüni P, Bartlett C. The value of patient flow charts in reports of randomized controlled trials: bibliographic study. *JAMA*. 2001;285:1996–9.
- Eng J. Sample size estimation: how many individuals should be studied? *Radiology*. 2003;227:309–13.
- Escribano M, Herrera D, Morante S, Teughels W, Quirynen M, Sanz M. Efficacy of a low-concentration chlorhexidine mouth rinse in non-compliant periodontitis patients attending a supportive periodontal care programme: a randomized clinical trial. *J Clin Periodontol*. 2010;37:266–75.
- Esposito M, Coulthard P, Worthington HV, Jøstad A. Quality assessment of randomized controlled trials of oral implants. *Int J Oral Maxillofac Implants*. 2001;16:783–92.
- Fenwick J, Needleman IG, Moles DR. The effect of bias on the magnitude of clinical outcomes in periodontology: a pilot study. *J Clin Periodontol*. 2008;35:775–82.
- Fergusson D, Glass KC, Waring D, Shapiro S. Turning a blind eye: the success of blinding reported in a random

- sample of randomised, placebo controlled trials. *BMJ*. 2004;328:432.
- Finch C, Braham R, McIntosh A, McCory P, Wolfe R. Should football players wear custom fitted mouthguards? Results from a group randomised controlled trial. *Inj Prev*. 2005;11:242–6.
- Fine D. Periodontal disease in Barnett ML, Pihlstrom BL. Methods for enhancing the efficiency of dental/oral health clinical trials: current status, future possibilities. *J Dent Res*. 2004;83:744–50.
- Fisher CG, Wood KB. Introduction to and techniques of evidence-based medicine. *Spine (Phila Pa 1976)*. 2007;32:S66–72.
- Fleiss JL. General design issues in efficacy, equivalency and superiority trials. *J Periodont Res*. 1992;27:306–13.
- Florey CD. Sample size for beginners. *BMJ*. 1993;306(6886):1181–4.
- Forder PM, Gebiski VJ, Keech AC. Allocation concealment and blinding: when ignorance is bliss. *Med J Aust*. 2005;182:87–9.
- Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317:141–5.
- Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C, CONSORT Group. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med*. 2006;144:364–7.
- Gansky SA, Ellison JA, Rudy D, Bergert N, Letendre MA, Nelson L, et al. Cluster-randomized controller trial of an athletic trainer-directed spit (smokeless) tobacco intervention for collegiate baseball athletes: results after 1 year. *J Athl Train*. 2005;40:76–87.
- Giraudeau B, Ravaud P, Donner A. Sample size calculation for cluster randomized cross-over trials. *Stat Med*. 2008;27:5578–85.
- GlaxoSmithKline. Clinical Trial Registry. <http://www.gsk-clinicalstudyregister.com/>. Accessed 11 May 2011.
- CONSORT Glossary 2007. Available at <http://www.consort-statement.org/resources/glossary/>. Accessed 10 May 2011.
- Gordon JS, Andrews JA, Lichtenstein E, Severson HH, Akers L. Disseminating a smokeless tobacco cessation intervention model to dental hygienists: a randomized comparison of personalized instruction and self-study methods. *Health Psychol*. 2005;24:447–55.
- Greene WL, Concato J, Feinstein AR. Claims of equivalence in medical research: are they supported by the evidence? *Ann Intern Med*. 2000;132:715–22.
- Greenstein G, Nunn ME. A method to enhance determining the clinical relevance of periodontal research data: number needed to treat (NNT). *J Periodontol*. 2004;75:620–4.
- Grimes DA, Schulz KF. Clinical research in obstetrics and gynecology: a Baedeker for busy clinicians. *Obstet Gynecol Surv*. 2002;57:S35–53.
- Gunsolley JC. Clinical efficacy of antimicrobial mouthrinses. *J Dent*. 2010;38:S6–10.
- Gunsolley JC, Elswick RK, Davenport JM. Equivalence and superiority testing in regeneration clinical trials. *J Periodontol*. 1998;69:521–7.
- Haahr MT, Hróbjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials*. 2006;3:360–5.
- Harrison JE. Clinical trials in orthodontics II: assessment of the quality of reporting of clinical trials published in three orthodontic journals between 1989 and 1998. *J Orthod*. 2003;30:309–15.
- Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol*. 1999;28:319–26.
- Heritier SR, Gebiski VJ, Keech AC. Inclusion of patients in clinical trial analysis: the intention-to-treat principle. *Med J Aust*. 2003;179:438–40.
- Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Intervention*. http://www.igh.org/Cochrane/tools/Ch16_Specialstatistics.pdf. Accessed 10 May 2011.
- Hodges JS. Intention-to-treat has implications for study planning and execution, not just subject retention and follow-up. *J Clin Periodontol*. 2008;35:683–4.
- Hopewell S, Altman DG, Moher D, Schulz KF. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal ‘Instructions to Authors’. *Trials*. 2008a;9:20.
- Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med*. 2008b;5:e20.
- Hróbjartsson A, Forfang E, Haahr MT, Als-Nielsen B, Brorson S. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *Int J Epidemiol*. 2007;36:654–63.
- Hróbjartsson A, Pildal J, Chan AW, Haahr MT, Altman DG, Gøtzsche PC. Reporting on blinding in trial protocols and corresponding publications was often inadequate but rarely contradictory. *J Clin Epidemiol*. 2009;62:967–73.
- Hugh M, Altman DG, Hammerschlag R, Hammerschlag R, Li YP, Wu TX, White A, Moher D, STRICTA Revision Group. Revised STAndards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement (Chinese version). *Zhong Xi Yi Jie He Xue Bao*. 2010;8:804–18.
- Hujoel PP, DeRouen TA. Determination and selection of the optimum number of sites and patients for clinical studies. *J Dent Res*. 1992;71:1516–21.
- International Clinical Trials Registry Platform. <http://www.who.int/ictpr/en/>. Accessed 11 May 2011.
- International Conference on Harmonisation (ICH). Choice of control group in clinical trials, ICH topic E10. London: EMEA; 2000.
- Ioannidis JP, Evans SJ, Gotzsche PC, O’Neill RT, Altman DG, Schulz K, et al.; CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*. 2004;141:781–8.
- Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1–12.

- James KE, Lee KK, Kraemer HC, Fuller RK. An index for assessing blindness in a multicenter clinical trial: disulfiram for alcohol cessation – a VA cooperative study. *Stat Med*. 1990;15:1421–34.
- Jeffcoat M. What is clinical significance? *J Clin Periodontol*. 2002;29:S30–2.
- Jeffcoat M, Parry S, Sammel M, Clothier B, Catlin A, Macones G. Periodontal infection and preterm birth: successful periodontal therapy reduces the risk of preterm birth. *BJOG*. 2011;118:250–6.
- Jokstad A, Esposito M, Coulthard P, et al. The reporting of randomized controlled trials in prosthodontics. *Int J Prosthodont*. 2002;15:230–42.
- Jönsson B, Öhrn K, Oscarson N, Lindberg P. The effectiveness of an individually tailored oral health educational programme on oral hygiene behaviour in patients with periodontal disease: a blinded randomized-controlled clinical trial (one-year follow-up). *J Clin Periodontol*. 2009;36:1025–34.
- Jönsson B, Öhrn K, Lindberg P, Oscarson N. Evaluation of an individually tailored oral health educational programme on periodontal health. *J Clin Periodontol*. 2010;37(10):912–9.
- Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42–6.
- Karlsson J, Engebretsen L, Dainty K, ISAKOS Scientific Committee. Considerations on sample size and power calculations in randomized clinical trials. *Arthroscopy*. 2003;19:997–9.
- Kelley K, Maxwell SE, Rausch JR. Obtaining power or obtaining precision. Delineating methods of sample-size planning. *Eval Health Prof*. 2003;26:258–87.
- Kelly PJ, Webster AC, Craig JC. How many patients do we need for a clinical trial? Demystifying sample size calculations. *Nephrology (Carlton)*. 2010;15:725–31.
- Kiang H, Bian Z, Tai BJ, Du MQ, Peng B. The effect of a bi-annual professional application of APF foam on dental caries increment in primary teeth: 24-month clinical trial. *J Dent Res*. 2005;84:265–8.
- Kim H-Y, Preisser JS, Rozier RG, Valiyaparambil JV. Multilevel analysis of group-randomized trials with binary outcomes. *Community Dent Oral Epidemiol*. 2006;34:241–51.
- Kingman A. Specific statistical consideration relevant to the design and analysis of gingivitis trials demonstrating product superiority or equivalence. *J Periodont Res*. 1992;27:378–89.
- Kirby A, Gebiski V, Keech AC. Determining the sample size in a clinical trial. *Med J Aust*. 2002;177:256–7.
- Koch GG, Paquette DW. Design principles and statistical considerations in periodontal clinical trials. *Ann Periodontol*. 1997;2:42–63.
- Kolahi J, Abrishami M. Multiple-blind: towards a new blinding protocol for future generations of clinical trials. *Med Hypotheses*. 2009;73:843–5.
- Kolahi J, Bang H, Park J. Towards a proposal for assessment of blinding success in clinical trials: up-to-date review. *Community Dent Oral Epidemiol*. 2009;37:477–84.
- Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev*. 2007;(2):MR000012.
- Lao L, Bergman S, Hamilton GR, Langenberg P, Berman B. Evaluation of acupuncture for pain control after oral surgery: a placebo-controlled trial. *Arch Otolaryngol Head Neck Surg*. 1999;125:567–72.
- LaRosa JC, Applegate W, Crouse 3rd JR, Hunninghake DB, Grimm R, Knopp R, et al. Cholesterol lowering in the elderly. Results of the cholesterol reduction in seniors program (crisp) pilot study. *Arch Intern Med*. 1994;154:529–39.
- Levine RJ, Carpenter WT, Appelbaum PS. Clarifying standards for using placebos. *Science*. 2003;300:1659–61.
- Lewis JA, Jönsson B, Kreutz G, Sampaio C, van Zwieten-Boot B. Placebo-controlled trials and the Declaration of Helsinki. *Lancet*. 2002;359(9314):1337–40.
- Listgarten MA. General issues in efficacy, equivalency, and superiority trials: clinical considerations. *J Periodont Res*. 1992;27:314–9.
- MacPherson H, Altman DG, Hammerschlag R, Li Y, Wu T, White A, Moher D, STRICTA Revision Group. Revised STAndards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. *Acupunct Med*. 2010a;28:83–93.
- MacPherson H, Altman DG, Hammerschlag R, Youping L, Taixiang W, White A, Moher D, STRICTA Revision Group. Revised STAndards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. *J Evid Based Med*. 2010b;3:140–55.
- Marshman Z, Farid F. The quality of reporting of randomised controlled trials in dental public health. *Community Dent Health*. 2010;27:253–6.
- Masouredis CM, Hilton JF, Grady D, Gee L, Chesney M, Hengl L, et al. A spit tobacco cessation intervention for college athletes: three-month results. *Adv Dent Res*. 1997;11:354–9.
- McAlister FA, Sackett DL. Active-control equivalence trials and antihypertensive agents. *Am J Med*. 2001;111:553–8.
- McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. *JAMA*. 1990;263:1371–6.
- Meinert CL. Clinical trials, design, conduct and analysis. New York: Oxford University Press; 1985. p. 71–89, 119–37, 288–96.
- Michalowicz BS, Hodges JS, DiAngelis AJ, Lupo VR, Novak MJ, Ferguson JE, et al. Treatment of periodontal disease and the risk of preterm birth. *N Engl J Med*. 2006;355:1885–94.
- Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials*. 1995;16:62–73.
- Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998;352(9128):609–13.

- Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before and after evaluation? *JAMA*. 2001a;285:1992–5.
- Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2001b;357(9263):1191–4.
- Montenegro R, Needleman I, Moles D, Tonetti M. Quality of RCTs in periodontology – a systematic review. *J Dent Res*. 2002;81:866–70.
- Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark: the reporting of blinding status in randomized controlled trials. *J Clin Epidemiol*. 2002;55:787–90.
- Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, Lacchetti C, Leung TW, Darling E, Bryant DM, Bucher HC, Schünemann HJ, Meade MO, Cook DJ, Erwin PJ, Sood A, Sood R, Lo B, Thompson CA, Zhou Q, Mills E, Guyatt GH. Randomized trials stopped early for benefit: a systematic review. *JAMA*. 2005;294:2203–9.
- Murray DM, Varnell SP, Blistein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. *Am J Public Health*. 2004;94:423–32.
- Neaton JD, Mugglin AS. From observational studies to randomized trials: asking the right question at the right time. *J Clin Periodontol*. 2006;33:517–9.
- Needleman I, Suvan J, Gilthorpe MS, Tucker R, St George G, Giannobile W, Tonetti M, Jarvis M. A randomized-controlled trial of low-dose doxycycline for periodontitis in smokers. *J Clin Periodontol*. 2007;34:325–33.
- Needleman I, Moher D, Altman DG, Schulz KF, Moles DR, Worthington H. Improving the clarity and transparency of reporting health research: a shared obligation and responsibility. *J Dent Res*. 2008;87:894–5.
- Newnham JP, Newnham IA, Ball CM, Wright M, Pennell CE, Swain J, et al. Treatment of periodontal disease during pregnancy: a randomized controlled trial. *Obstet Gynecol*. 2009;114:1239–48.
- Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: basic principles and common pitfalls. *Nephrol Dial Transplant*. 2010;25:1388–93.
- Noordzij M, Dekker FW, Zoccali C, Jager KJ. Sample size calculations. *Nephron Clin Pract*. 2011;118:c319–23.
- Odgaard-Jensen J, Vist GE, Timmer A, Kunz R, Akl EA, Schünemann H, Briel M, Nordmann AJ, Pregno S, Oxman AD. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev*. 2011;(4):MR000012.
- Pandis N, Polychronopoulou A, Eliades T. An assessment of quality characteristics of randomised control trials published in dental journals. *J Dent*. 2010;38:713–21.
- Park J, Bang H, Cañette I. Blinding in clinical trials, time to do it better. *Complement Ther Med*. 2008;16:121–3.
- Persson GR, Samuelsson E, Lindahl C, Renvert S. Mechanical non-surgical treatment of peri-implantitis: a single-blinded randomized longitudinal clinical study. II. Microbiological results. *J Clin Periodontol*. 2010;37:563–73.
- Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ, CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA*. 2006;295:1152–60.
- Pihlstrom BL. Reporting clinical trial results. *J Am Dent Assoc*. 2009;140:14–5.
- Pihlstrom BL, Barnett ML. Design, operation, and interpretation of clinical trials. *J Dent Res*. 2010;89:759–72.
- Pildal J, Hróbjartsson A, Jørgensen KJ, Hilden J, Altman DG, Gøtzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol*. 2007;36(4):847–57.
- Rees JR, Wade TJ, Levy DA, Colford Jr JM, Hilton JF. Changes in beliefs identify unblinding in randomized controlled trials: a method to meet consort guidelines. *Contemp Clin Trials*. 2005;26:25–37.
- Rid A, Schmidt H. The 2008 Declaration of Helsinki – first among equals in research ethics? *J Law Med Ethics*. 2010;38:143–8.
- Scales DC, Rubenfeld GD. Estimating sample size in critical care clinical trials. *J Crit Care*. 2005;20:6–11.
- Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet*. 2002a;359(9306):614–8.
- Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet*. 2002b; 359(9307): 696–700.
- Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet*. 2002c;359(9305):515–9.
- Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet*. 2002d;359(9308):781–5.
- Schulz KF, Grimes DA. Unequal group sizes in randomised trials: guarding against guessing. *Lancet*. 2002e;359(9310):966–70.
- Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet*. 2005;365(9467):1348–53.
- Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
- Simpson JM, Klar N, Donner A. Accounting for cluster randomization: a review of primary prevention trials, 1990 through 1993. *Am J Public Health*. 1995;85: 1378–83.
- Sjögren P, Halling A. Quality of reporting randomised clinical trials in dental and medical research. *Br Dent J*. 2002;192:100–3.
- Stoner JA, Payne JB. Interpretation of treatment effects in periodontal research: a note on the number needed to treat. *J Can Dent Assoc*. 2008;74:435–7.
- Turpin DL. CONSORT and QUOROM guidelines for reporting randomized clinical trials and systematic reviews. *Am J Orthod Dentofacial Orthop*. 2005;128(6):681–5.

- Tu YK, Blance A, Clerehugh V, Gilthorpe MS. Statistical power for analyses of changes in randomized controlled trials. *J Dent Res*. 2005;84:283–7.
- Tu Y-K, Maddick I, Kellett M, Clerehugh V, Gilthorpe MS. Evaluating the quality of active-control trials in periodontal research. *J Clin Periodontol*. 2006;33:151–6.
- van der Putten GJ, De Visschere L, Schols J, de Baat C, Vanobbergen J. Supervised versus non-supervised implementation of an oral health care guideline in (residential) care homes: a cluster randomized controlled clinical trial. *BMC Oral Health*. 2010;10:17.
- Van der Weijden F, Slot DE, Rosema M, Timmerman M. ITT in respect to GCP: a matter of diligence. *J Clin Periodontol*. 2008;35:681–2.
- Varela VM, Heller D, Silva-Senem MX, Torres MC, Colombo AP, Feres-Filho EJ. Systemic antimicrobials adjunctive to a repeated mechanical and antiseptic therapy for aggressive periodontitis: a 6-month randomized controlled trial. *J Periodontol*. 2011;82:1121–30.
- Varnell SP, Murray DM, Janega JB, Blitstein JL. Design and analysis of group-randomized trials: a review of recent practices. *Am J Public Health*. 2004;94:393–9.
- Walter SD, Awasthi S, Jeyaseelan L. Pre-trial evaluation of the potential for unblinding in drug trials: a prototype example. *Contemp Clin Trials*. 2005;26:459–68.
- Wei LJ, Lachin JM. Properties of the urn randomization in clinical trials. *Control Clin Trials*. 1988;9:345–64.
- Weingarten MA, Paul M, Leibovici L. Assessing ethics of trials in systematic reviews. *BMJ*. 2004;328(7446):1013–4.
- World Health Organization Website. http://www.who.int/ictrp/data_set/en/index1.html. Accessed 11 May 2011.
- Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D, CONSORT Group, Pragmatic Trials in Healthcare (Practihc) Group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. 2008;337:a2390.

Bias is the systematic deviation of inferences or results from the truth, or processes that bring about such deviation (Dunn et al. 2003). Biases can operate in either direction: Different biases can lead to underestimation or overestimation of the true intervention effect. Biases can vary in magnitude: some are small (and trivial compared with the observed effect) and some are substantial (so that an apparent finding may be entirely due to bias). Because the results of a study may in fact be unbiased despite a methodological flaw, it is more appropriate to consider **risk of bias** (Higgins and Green 2011).

Bias should not be confused with **imprecision**. Bias refers to *systematic error*, meaning that multiple replications of the same study would reach the wrong answer on average. Imprecision refers to *random error*, meaning that multiple replications of the same study will produce different effect estimates because of sampling variation even if they would give the right answer on average. The results of smaller studies are subject to greater sampling variation and hence are less precise. Imprecision is reflected in the confidence interval around the intervention effect estimate from each study and in the weight given to the results of each study in a meta-analysis. More precise results are given more weight (Higgins and Green 2011).

Biases associated with particular characteristics of studies may be examined using a technique often known as **meta-epidemiology**. A meta-epidemiological study analyses a collection of meta-analyses, in each of which the component

studies have been classified according to some study-level characteristic. A simple analysis of a meta-epidemiological study is to calculate the ‘ratio of odds ratios’ within each meta-analysis (e.g. the intervention odds ratio in trials with inadequate/unclear allocation concealment divided by the odds ratio in trials with adequate allocation concealment). These ratios of odds ratios are then combined across meta-analyses, in a meta-analysis. Thus, such analyses are also known as ‘meta-meta-analyses’ (Higgins and Green 2011).

Bias can occur at any phase of research, including study design or data collection, and in the process of data analysis and publication (**Fig. 5.1**). Numerous type of bias have been described, some of which are described below. Bias is not a dichotomous variable. Interpretation of bias cannot be limited to a simple inquisition: Is bias present or not? Instead, reviewers of the literature must consider the degree to which bias was prevented by proper study design and implementation. **Table 5.1** provides a summary of different types of bias, when they occur, and how they might be avoided (Pannucci and Wilkins 2010).

5.1 Bias Before a Trial Begins

5.1.1 Susceptibility Biases

Selection criteria are necessary to minimize group variability so that statistically valid conclusions can be reached in a reasonably sized sample.

Fig. 5.1 Major sources of bias in clinical research (Pannucci and Wilkins 2010. Reprinted with permission from Wolters Kluwer Health)

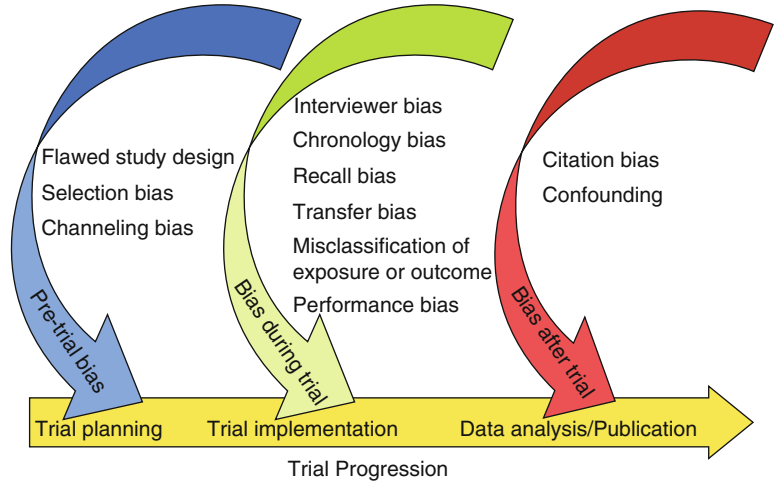


Table 5.1 Tips to avoid different types of bias during a trial (Pannucci and Wilkins 2010) (Reprinted with permission from Wolters Kluwer Health)

Type of bias	How to avoid
Pre-trial bias	
Flawed study design	• Clearly define risk and outcome, preferably with objective or validated methods. Standardize and blind data collection.
Selection bias	• Select patients using rigorous criteria to avoid confounding results. Patients should originate from the same general population. Well designed, prospective studies help to avoid selection bias as outcome is unknown at time of enrollment.
Channeling bias	• Assign patients to study cohorts using rigorous criteria.
Bias during trial	
Interviewer bias	• Standardize interviewer’s interaction with patient. Blind interviewer to exposure status.
Chronology bias	• Prospective studies can eliminate chronology bias. Avoid using historic controls (confounding by secular trends).
Recall bias	• Use objective data sources whenever possible. When using subjective data sources, corroborate with medical record. Conduct prospective studies because outcome is unknown at time of patient enrollment.
Transfer bias	• Carefully design plan for lost-to-follow-up patients prior to the study.
Exposure Misclassification	• Clearly define exposure prior to study. Avoid using proxies of exposure.
Outcome Misclassification	• Use objective diagnostic studies or validated measures as primary outcome.
Performance bias	• Consider cluster stratification to minimize variability in surgical technique.
Bias after trial	
Citation bias	• Register trial with an accepted clinical trials registry. Check registries for similar unpublished or in-progress trials prior to publication.
Confounding	• Known confounders can be controlled with study design (case control design or randomization) or during data analysis (regression). Unknown confounders can only be controlled with randomization.

But if selection criteria have prognostic importance, they create a type of susceptibility bias called **selection bias** (Paradis 2008). The common consequence of selection bias is that the association between exposure and outcome among those selected for analysis differs from

the association among those eligible (Hernàn et al. 2004).

Lopez et al. (2007) evaluated the occurrence of sources of selection bias in case–control studies published in English during the year 2004. In relatively few studies did the authors provided

information on recruitment periods for cases and controls (31.1% and 20%, respectively), sampling methods (26.7% and 31.1%, respectively), or participation rates (8.9% and 6.7%, respectively). The source of control subjects was appropriate in 15.6% of the studies, and the strategy used to select the controls was adequate in only 8.9% of the studies.

It has been showed that that several types of bias may be present in the selection of study participants in a considerably large number of publications on periodontitis (Lopez et al. 2007). The optimal control group for a case–control study is randomly sampled from the same population that gave rise to the cases during a clearly specified time interval. Still, the list of inadequate sources for control subjects found in the biomedical literature is long, ranging from friends of cases to staff members at dental schools. Lopez et al. (2007) showed that a particularly large group of control subjects in periodontal research originate from health-care institutions. Though common, this practice is problematic and is likely to introduce several types of selection bias, such as referral and admission biases.

Lopez et al.'s (2007) findings suggest that membership bias is also frequently introduced in periodontal research when members of staff, students, mixed groups, and/or other groups in a health-care institution constitute the control group. This type of control subjects is likely to differ systematically in many aspects from the cases, for example, with reference to age, smoking, and oral hygiene habits. They do not originate from the source population of the cases and should thus not be used as controls. An additional argument to avoid institution-based case–control studies on periodontitis is that they can introduce a type of selection bias known as prevalence incidence bias because the cases referred to a periodontal clinic are not necessarily representative of the cases found in the community (Lopez et al. 2007).

Channelling bias is a type of selection bias often seen in observational studies comparing older versus newer drugs from similar therapeutic classes (Lobo et al. 2006; Petri and Urquhart 1991). Channeling occurs when drug therapies with similar indications, either self-selected or clinically assigned, are prescribed to groups of patients with varying baseline prognoses. Drugs with similar

therapeutic actions may be launched into the market at different times. This creates a situation where drugs entering the market at a later stage may have a higher likelihood of being prescribed to patients who have not responded well to existing agents. The promotional claims for the newer drug often conjure up the profile of a patient most suitable and more likely to benefit from the newer therapy, leading to channeling. Thus, recipients of the older drug and newer drug will differ in terms of their baseline morbidities. Hence, channeling ultimately leads to the selective prescription of drugs from the same therapeutic class to groups of patients differing in their susceptibility to problems or special pre-existing morbidities (Lobo et al. 2006).

At the launch of a newer drug, manufacturers seek to differentiate their product from the competition generally in terms of better efficacy claims and/or side effect profiles, if any. Promotional claims of a better side effect profile might lead physicians to prescribe the new drug for patients who have experienced undesirable side effect(s) from the use of existing drugs. In case of drugs claiming better efficacy, physicians might prescribe them to patients who have failed on prior therapies. Thus, the therapeutic agent entering the market at a later stage often ends up being selectively prescribed to patients with a more severe clinical profile (Lobo et al. 2006).

Studies using a historical comparator group often have **chronology bias**, a bias that occurs because of changes over time, which affect outcome. Such changes can occur in the way treatments are delivered, diseases are detected, or even in the methods used to measure variables or outcome (Paradis 2008). There are many subtle forms of chronology bias, but the most flagrant is the use of historical controls. It is clear that diagnostic and therapeutic ability may change over time. In addition, supportive therapy not directly related to the treatment under study may improve, thus influencing the outcome (Haines 1979).

5.1.2 Design Biases

Before the first subject is enrolled, various *design features* beyond subject susceptibility can establish an environment that facilitates or deters bias.

Flexibility in design, such as vague definitions of outcome or analysis methods, encourages bias by enabling investigators to manipulate a study to obtain an answer they want or think they should get instead of objectively unearthing it (Paradis 2008).

The evaluation of periodontitis interventions presents several challenges due to the disease's heterogeneity and its irregular, episodic pattern; nevertheless, the intent of these novel interventions is to prevent, diagnose, inhibit, or reverse periodontal disease progression. Careful consideration of the trial's objectives should dictate clinical endpoints (primary and surrogate), comparison groups (placebo, standard therapy, test therapy), and equivalence versus superiority as the basis for conclusions. Several design elements such as control population specification, randomization, masking, sample size calculation, and standardization of procedures for patient care and assessment can decrease potential bias and variability. In both parallel and paired (split-mouth) design trials, multiplicities of endpoints, treatments, and subgroups require strategies which address the broader scope of chance findings without excessive loss of study power. Also, the longitudinal assessment of multiple periodontal sites within patients produces correlated data structures for which analytic methods need to account for the appropriate sampling unit (Koch and Paquette 1997).

A trial design factor that can contribute to bias is the use of **surrogate endpoints**. Endpoints are conditions or events that are associated with individual study subjects and that are used to assess treatment efficacy. Two types of endpoints can be distinguished: 'true' endpoints (reflect unequivocal evidence of tangible benefit to the patient) and 'surrogate' endpoints (usually a measure of disease process) (Hujoel and DeRouen 1995).

True endpoints are outcomes that directly measure how a patient feels, functions, or survives. **True endpoints** are tangible to the patient. The word 'tangible' is defined by the dictionary as 'capable of being precisely identified or realized by the mind'. Tooth loss or painful periodontal abscesses are examples of events that can precisely be identified or realized by the patient's mind, that is, that are tangible. True endpoints also include subjective oral health-related

quality-of-life measurements or simple self-reported symptoms such as bleeding after brushing. True endpoints are sometimes referred to as clinically relevant endpoints, clinically meaningful endpoints, terminal endpoints, or ultimate endpoints (Hujoel 2004).

For periodontics, the prevention of edentulism or tooth loss would be such a criterion of a true endpoint; however, tooth loss may have a rate of occurrence which is too small to justify its consideration as an endpoint in terms of feasible duration of follow-up. Clinical trials attempting to measure periodontal treatment responses in terms of tooth loss rates would therefore require exceedingly large numbers of patients followed for long assessment periods (Koch and Paquette 1997). Furthermore, the cause of tooth loss is not always easy to definitively determine. Because histologic examination and prevention of tooth loss after treatment are not realistic endpoint measures for large-scale, multicenter periodontal therapy studies, clinical researchers have selected surrogate endpoints, or changes in patient status, which are considered to be strong predictors of a true endpoint (Lynch et al. 2006).

Surrogate endpoints are intangible outcomes used as a substitute for a true endpoint. Surrogate endpoints are intangible to the patient's mind (Hujoel 2004). Surrogate endpoints save costs and time. Their disadvantages can be the uncertainty about their correlation to significant clinical outcomes, their inability to give an accurate picture of complications, especially long-term ones, and their inability, because they measure short-term changes, to address long-term effects such as implant longevity and the variable ways in which they are measured (Paradis 2008). Because tooth loss is not a realistic response criterion under these assumptions, changes in patient status which are considered to be strong predictors of tooth loss, or surrogate endpoints, are the next best choice. Two reasonable candidates for surrogate endpoints for tooth loss are changes in probing attachment level and radiographic alveolar bone (height or mass) since substantial changes in either could essentially lead to tooth loss (Koch and Paquette 1997). Laboratory determinations of biological activity (e.g. microbiology, immunology) at periodontal sites are potential diagnostic and prog-

nostic tests currently being considered (Koch and Paquette 1997; Loos and Tjoa 2005).

To assess the cumulative beneficial effects of treatment modalities in medicine, clinical researchers in other disciplines have employed the use of **composite endpoint** analysis. These endpoints are traditionally expressed as binary outcomes classified as success or failure at the patient level. The endpoints are then amenable to traditional statistical analysis (Fisher exact test, χ^2 test, and logistic regression) to establish statistical significance. Recently, this approach was successfully used to support FDA (e.g. U.S. Food and Drug Administration) approval for a periodontal regeneration product. Composite endpoints included in the analyses for this clinical trial combined the clinical attachment level parameter (DCAL) with radiological parameters (percent bone fill: %BF; or linear bone gain: LBG), offering a potential improvement over single-outcome measures for determining clinically meaningful effects on both new bone formation and attachment levels. The argument for such an approach comes from the fact that CALs measure the changes in soft tissue but do not directly assess whether there has been any new supporting tissues (i.e. bone, periodontal ligament, and cementum) formation, an objective of periodontal regenerative therapies. Alternatively, radiographic analysis and surgical re-entry procedures have the potential to determine whether new tissue (i.e. bone) has been formed but do not measure changes in attachment of the tooth to the supporting bone (Lynch et al. 2006).

The construction of a composite endpoint involves the selection of clinically meaningful thresholds for each component. In this situation, previously approved FDA standards provided thresholds. Lynch et al. (2006) used the higher standards for DCAL (2.67 mm), mean radiographic change in LBG (1.1 mm), and mean radiographic change in %BF (14.1%). These measurements represent data submitted to federal regulatory agencies for previously approved products and provide data for each parameter assessed radiographically (not clinically, i.e. surgical re-entry). Criteria were selected to achieve success–fail outcomes where thresholds for components correspond to clinically meaningful improvements from baseline (Lynch et al. 2006).

Hujoel and DeRouen (1995) performed a survey on aspects of endpoint usage in randomized controlled trials (RCTs) on the treatment of periodontitis. Ninety-two publications (1988–1992) reporting on 82 RCTs were identified. The typical number of endpoints per RCT was 6 (range: 1–28). The three most frequently used endpoints were mean probing depth (78% of the trials), mean probing attachment level (66%), and the plaque index (37%). In total, 153 distinct surrogate endpoints were defined. Most of these were used infrequently; over 80% of the 153 endpoints were used in fewer than 5 of the 82 trials. No trials used tooth loss as a true endpoint. Surrogate endpoints based on re-entry surgery were exclusively used for regenerative procedures, and microbiological surrogate endpoints were mostly used for RCTs on antimicrobials.

More recently, two comprehensive literature reviews, evaluating endpoint measurements from clinical trials evaluating growth and amelogenin-like factors (Giannobile and Somerman 2003) or bone replacement grafts (Reynolds et al. 2003) for the treatment of periodontal osseous defects, have provided perspectives on benchmarks of success (Lynch et al. 2006).

In the first systematic review, Giannobile and Somerman (2003) evaluated eight studies, representing seven RCTs and one quasi-experimental study, representing a total population of 511 subjects who were analysed with respect to Emdogain. The common clinical endpoints in these studies were clinical attachment level, probing depth, or bone level (radiographic, re-entry, or histologic).

In the second review, the therapeutic endpoints examined for assessing the efficacy of bone replacement grafts in the treatment of periodontal osseous defects included changes in bone level, clinical attachment level, probing depth, gingival recession, and crestal resorption. For purposes of meta-analysis, change in bone level (bone fill) was used as the primary outcome measure, measured upon surgical re-entry or transgingival probing (sounding) (Reynolds et al. 2003).

Hujoel (2004) recently concluded that as the goal of clinical research is to provide unequivocal evidence regarding the potential tangible benefits of a treatment, periodontal research cannot afford to keep stopping short of this goal. Two steps

need to be taken. First, research needs to be initiated to evaluate which periodontal surrogates are most reliable. Establishing such an evidence-based assessment, rather than a notion-based assessment, of periodontal surrogates will provide a scientific foundation for the conduct of exploratory trials. Second, true endpoint pivotal trials need to be initiated to provide unequivocal evidence that the promising treatment identified by means of surrogates truly provides tangible patient benefits. These two steps will provide the necessary evidence to put clinical research on a firmer scientific footing (Hujoel 2004).

In studies of peri-implantitis therapy, ‘true’ endpoints are preferred to ‘surrogate’, because

they can capture a substantial proportion of the effect of treatment on an outcome of interest, for example, implant failure. However, it has revealed that most endpoints in studies of peri-implantitis therapy are surrogate markers for important clinical events such as implant failure. In fact, no study employed any true endpoint as a primary measure of outcome, because implant failure was only reported as a consequence of treatment. Mean PPD, mean BOP, and mean CAL at implant were the most frequently reported clinical surrogate endpoints in the studies selected (Table 5.2). These endpoints have not yet been validated as surrogates for true clinical endpoints, for example, implant failure, however (Faggion et al. 2010).

Table 5.2 Types of clinical endpoints and definitions used in studies on peri-implantitis therapy The third column gives the frequency of clinical endpoints across the studies (Faggion et al. 2010) (Reprinted with permission from Elsevier)

Endpoint	Definition	Frequency
Full-mouth plaque score	Presence/absence of dental plaque along the gingival/mucosal margin following the use, or not, of disclosing dye and expressed as a percentage of the sites examined for each subject (four or six sites per tooth and implant)	3
Local plaque score	Presence of dental plaque along the mucosal margin at four sites of the treated implant, recorded after the use, or not, of a disclosing dye	3
PPD at worst site	Pocket probing depth at the deepest site at implant	2
Mean PPD at implant	Distance (mm) between the gingival/mucosal margin and the bottom of the sulcus/pocket	12
Bleeding at implant (without PD)	Bleeding resulting from running a probe along the soft tissue margin without probe penetration inside the sulcus or pocket	2
BOP after PD at implant	Presence or absence of bleeding for up to 15 or 30 s after gentle probing	7
Full-mouth MB/GB	Bleeding resulting from running a probe along the soft tissue margin without probe penetration inside the sulcus or pocket (teeth and implants)	1
Full-mouth BOP	Presence or absence of bleeding for up to 15 s after gentle probing at teeth and implants	3
Suppuration	Presence/absence of spontaneous suppuration or suppuration on probing	3
Full-mouth PD	Distance (mm) between the gingival/mucosal margin and the bottom of the sulcus/pocket at implants and teeth	1
CAL at implant	Distance between the cemento-enamel junction and the bottom of the tooth sulcus/pocket for all teeth in partially edentulous subjects (measured at six sites)	8
CAL at teeth	Distance (mm) between the cemento-enamel junction and the bottom of the tooth sulcus/pocket measured at four to six sites	1
BOS	Bleeding after sampling for microbiological assessment	1
PI or modified PI (mPI) at implant	Determined on the mesial, facial, distal, and oral surfaces of the implant (score 0=no detection of plaque, score 1=plaque only recognised by running a probe across the smooth marginal surface of the implant, score 2=plaque can be seen by the naked eye, score 3=abundance of soft matter)	5

(continued)

Table 5.2 (continued)

Endpoint	Definition	Frequency
GR	Gingival recession measured from the implant neck (shoulder) to the mucosal margin (apical to the implant shoulder)	5
Supra-alveolar measurement	Unclear	1
Circumferential measurement	Length of the defect (mm) in the horizontal direction measured at the time of surgery	1
Intrabony measurement	Intrabony vertical defect (mm) measured at the time of surgery	1
Bleeding index or sulcus bleeding index	Assessed, after probing, at four to six sites (score 0=no bleeding when a periodontal probe is passed along the gingival margin adjacent to the implant, score 1=isolated bleeding spot visible, score 2=red line of blood on margin, score 3=heavy bleeding)	3
Full-mouth PI	Determined on the mesial, facial, distal, and oral surfaces of the implant and teeth (score 0=no detection of plaque, score 1=plaque only recognised by running a probe across the smooth marginal surface of the implant, score 2=plaque can be seen by the naked eye, score 3=abundance of soft matter)	1
DIM	Pseudopockets=linear distance from the free mucosal margin to the implant shoulder when the mucosal margin is coronal to the implant shoulder	1
DIB	Distance from the implant shoulder to the first bone contact	1
Probing bone level	Distance from the implant shoulder to the deepest depth at which the probe met strong resistance from contact with bone	1
Implant mobility	Horizontal implant movement tested manually	1
Others	Hyperplasia	1

5.2 Bias During Trial Execution

5.2.1 Information Biases

Information bias is the result of misclassification of study participants with respect to disease or exposure status. Thus, the concept of information bias refers to those people actually included in the study, whereas selection bias refers to the selection of the study participants from the source population, and confounding (see below) generally refers to non-comparability of subgroups within the source population. It is customary to consider two types of misclassification: non-differential and differential (Pearce et al. 2007).

Non-differential misclassification of exposure occurs when the probability of exposure misclassification is not related to disease status—that is, if diseased and non-diseased people are equally likely to be misclassified according to exposure. Similarly, misclassification of disease status is non-differential if exposed and non-exposed people are equally likely to be misclassified according to disease status. Differential misclassification occurs when the probability of misclassification of expo-

sure is different in diseased and non-diseased people, or the probability of misclassification of disease is different in exposed and non-exposed people (Pearce et al. 2007).

Information biases occur in both observational studies and RCTs. Two forms of information bias are detection bias and ascertainment bias (Paradis 2008).

Detection bias occurs if the methods for measuring outcome are not uniform either within a study or between a test and a comparator group. Vaguely defined outcomes, un-blinded evaluators and subjects, and unreliable surrogate endpoints facilitate it. Detection bias may also occur if outcomes are measured differently in a test and non-concurrent comparator group, especially if surrogate endpoints are not measured in exactly the same way in both groups. For reasons already discussed, this is especially likely to be a problem with historical controls. Similarly, blinding evaluators and/or subjects in one study but not in the other may bias outcome comparison. Subjects who know which treatment they received may differ from those who do not in how they report beneficial and/or harmful effects, and in their ten-

dency to seek treatment outside of a study and/or in their rate of dropping out of a study (Paradis 2008).

Establishing examiner reliability is a vital component of clinical periodontal research. Failure to adequately train and calibrate examiners in the collection of periodontal measures results in imprecise and unreliable data and jeopardizes the veracity of subsequent inference (Hill et al. 2006).

In distinction to detection bias, **ascertainment bias** results from a distorted determination of exposure to a factor that is important in a study. This can compromise the reliability of selection criteria or other factors related to outcome interpretation (Paradis 2008). For instance, one explanation for the lower prevalence of various forms of dementia in high-functioning elders is that factors such as higher education simply confound the diagnosis of dementia (Stern 2002). That is, markers of reserve (e.g. education, occupation, intelligence) may be confounded with measures of outcome (prevalence or incidence of Alzheimer's disease) (Storandt and Morris 2010). Recall bias, a type of ascertainment bias, can occur if pivotal information item is based on remembering previous events. Subjects with a disease tend to seek a reason for that disease, whereas with no motivation for such reflection, healthy individuals are less likely to recall such events. For example, subjects with rheumatoid arthritis are more likely to remember an ancestor with arthritis interpreted as rheumatoid arthritis, than are those not affected (Schull & Cobb 1969). Thus ascertainment bias has complicated the interpretation of association between periodontal infection and rheumatoid arthritis development in the First National Health and Nutrition Examination Survey and its epidemiological follow-up study (Demmer et al. 2011).

5.2.2 Transfer Bias

In almost all clinical studies, subjects are lost to follow-up. In these instances, investigators must consider whether these patients are fundamentally different than those retained in the study. Researchers must also consider how to treat patients lost to follow-up in their analysis. Well-

designed trials usually have protocols in place to attempt telephone or mail contact for patients who miss clinic appointments. Transfer bias can occur when study cohorts have unequal losses to follow-up. This is particularly relevant in surgical trials when study cohorts are expected to require different follow-up regimens. Some authors suggest that patient loss to follow-up can be minimized by offering convenient office hours, personalized patient contact via phone or email, and physician visits to the patient's home (Pannucci and Wilkins 2010).

5.2.3 Performance Bias

Performance bias occurs when procedures or interventions are not performed in a uniform way. It affects internal validity. Elimination of performance bias requires an intervention to have a fixed nature and uniform action (Paradis 2008).

As part of the part of the team and center training, the therapist who is administering an experimental periodontal treatment need a special training program. The clinical procedure protocol should also clearly define the technical endpoint criteria and conditions which are expected from the therapist, as well as any conditions relating to time of the procedure and use of anesthetic agents. These guidelines should be reviewed with the therapist and practiced in a clinical situation using a standardized set of instruments (Polson 1997).

5.2.4 Attrition Bias

The informed consent process explicitly guarantees each participant the right to freely terminate study participation at any time. This right is critically important and represents an ethical principle that helps with the protection of human research subjects. One undesirable, yet highly likely, cost of this protection is that, as study duration progresses, dropout, which can be thought of as another type of self-selection, compromises randomization. This has the potential to introduce attrition bias, cause imbalance among the previously randomized groups and,

thus, threaten the internal validity of the RCT (Leon et al. 2006; Flick 1988).

Attrition interferes with the objectives of an RCT as nonequivalent comparison groups can lead to biased estimates of the treatment effect. An attrition-induced reduction in sample size corresponds with a reduction in statistical power and a restricted sample that limits the generalizability of results (Leon et al. 2006).

The consequences of attrition include reduced statistical power and limitations on generalizability of the RCT results. An obvious problem with attrition is the missed assessments. This, of course, reduces the amount of data available for analyses and that corresponds to a reduction in statistical power. Furthermore, results from an RCT that has incomplete data can only be generalized to patients with characteristics of the participants who provided the data used in analyses. In other words, if there are demographic or clinical features that are associated with elevated dropout, the RCT will not inform clinical practice about patients with those features. The CONSORT (Consolidated Standards of Reporting Trials) Statement recommends that the flow of participants in an RCT be clearly documented in the publication that reports the RCT results (Moher et al. 2001a, b, c). This includes attrition rates for each group and the reasons that participants were excluded from the analyses (Leon et al. 2006).

When attrition does occur, three general approaches to analysis of incomplete data were suggested to illustrate the implication of these assumptions: (1) analysis of complete cases, (2) missing data imputation, and (3) analysis of incomplete data (Leon et al. 2006).

5.3 Bias After a Trial Ends

5.3.1 Citation Bias

Honest reporting begins with revealing the existence of all clinical studies, even those that reflect unfavourably on a research sponsor's product. Unfortunately, selective reporting of trials does occur, and it distorts the body of evidence available for clinical decision-making. Researchers (and journal editors) are generally most enthusi-

astic about the publication of trials that show either a large effect of a new treatment (positive trials) or equivalence of two approaches to treatment (non-inferiority trials). Researchers (and journals) typically are less excited about trials that show that a new treatment is inferior to standard treatment (negative trials) and even less interested in trials that are neither clearly positive nor clearly negative, since inconclusive trials will not in themselves change practice. This is **trial selection bias** (DeAngelis et al. 2004).

To address this problem, the International Committee of Medical Journal Editors (ICMJE) recommended that the editors should consider seriously for publication any carefully done study of an important question, relevant to their readers, whether the results for the primary or any additional outcome are statistically significant. Failure to submit or publish findings because of lack of statistical significance is an important cause of publication bias (www.icmje.org).

The ICMJE, in 2004, also required registration of all clinically directive trials, which it defined as 'any research project that prospectively assigns human subjects to intervention or comparison groups to study the cause-and-effect relationship between a medical intervention and a health outcome'. In May 2005, the ICMJE clarified this definition to exclude preliminary trials designed to study pharmacokinetics or major unknown toxicity (phase 1 trials). However, the ICMJE recognizes the potential benefit of having information about preliminary trials in the public domain, because these studies can guide future research or signal safety concerns. Consequently, in 2007, the ICMJE is expanding the definition of the types of trials that must be registered to include these preliminary trials and adopts the WHO's definition of clinical trial: 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes' and which begin enrollment after July 1, 2008. Health-related interventions include any intervention used to modify a biomedical or health-related outcome (e.g. drugs, surgical procedures, devices, behavioural treatments, dietary interventions, and process-of-care

changes). Health outcomes include any biomedical or health-related measures obtained in patients or participants, including pharmacokinetic measures and adverse events (Laine et al. 2007).

To assess a trial's external and internal validity, clinicians need information about a trial's design and conduct. Incomplete trial information makes this difficult. Many trial reports lack important information such as clearly defined endpoints and inclusion and exclusion criteria and are non-specific about randomization, blinding procedures, and/or sample size calculations. The ICMJE, by specifying a minimal data set that must be included when registering a trial, addresses some of these reporting problems. That data set is, however, very basic. For example, although it does include target sample size, it does not ask how that number was calculated (Paradis 2008). The original CONSORT (Consolidated Standards of Reporting Trials)-developed statement in 1996 and modified in 2001 has a more complete set of reporting criteria to help authors improve reporting of simple two-group parallel RCTs by using a checklist and flow diagram. The checklist items pertain to the content of the Title, Abstract, Introduction, Methods, Results, and Discussion. The revised checklist includes 22 items selected because empirical evidence indicates that not reporting the information is associated with biased estimates of treatment effect or because the information is essential to judge the reliability or relevance of the findings. The revised flow diagram depicts information from four stages of a trial (enrollment, intervention allocation, follow-up, and analysis). The diagram explicitly includes the number of participants, for each intervention group, that are included in the primary data analysis. Inclusion of these numbers allows the reader to judge whether the authors have performed an intention-to-treat analysis (Moher et al. 2001a, b, c).

In sum, the CONSORT statement is intended to improve the reporting of an RCT, enabling readers to understand a trial's conduct and to assess the validity of its results. However, CONSORT does not address other facets of reporting that also require attention, such as scientific content and readability of RCT reports, and also do not apply to observation trials that are

more prevalent (Moher et al. 2001a, b, c; Paradis et al. 2008). This was achieved by a consensus statement, STROBE, that stands for an international, collaborative initiative of epidemiologists, methodologists, statisticians, researchers, and journal editors involved in the conduct and dissemination of observational studies, with the common aim of **strengthening the reporting of observational studies in epidemiology**.

5.4 Overarching Biases

5.4.1 Optimism Bias

The belief that new therapies are better than established ones, optimism bias, influences both investigators and subjects. Such bias is based not on reality but on unrealistic hope. It can bias subjective evaluations of outcome as in the hypothetical knee motion example. And it can interfere with subjects' and investigators' equipoise, making it difficult to enrol subjects in trials comparing a new, but not fully investigated, therapy to conventional therapy (Paradis 2008).

5.4.2 Conflicts of Interest

Conflicts of interest have been defined as 'a set of conditions in which professional judgement concerning a primary interest (e.g. patient welfare, research validity) may be unduly influenced by a secondary interest (e.g. financial gain)'. Blinding, by minimizing selection, detection, and ascertainment biases, can help to prevent corruption of trial outcomes. Randomization can reduce selection bias. Finally, disclosure of conflicts of interest may allow those assessing trial results to consider whether they should examine that trial with reserve (Paradis 2008). To address this problem, ICMJE requests that when authors submit a manuscript, whether an article or a letter, they are responsible for disclosing all financial and personal relationships that might bias their work. To prevent ambiguity, authors must state explicitly whether potential conflicts do or do not exist. Authors should identify individuals who provide writing or other assistance

and disclose the funding source for this assistance. Investigators must disclose potential conflicts to study participants and should state in the manuscript whether they have done so. Authors should describe the role of the study sponsor, if any, in study design; collection, analysis, and interpretation of data; writing the report; and the decision to submit the report for publication. If the supporting source had no such involvement, the authors should so state. Biases potentially introduced when sponsors are directly involved in research are analogous to methodological biases. Some journals, therefore, choose to include information in the Methods section about the sponsor’s involvement (http://www.icmje.org/ethical_4conflicts.html).

Editors may request that authors of a study funded by an agency with a proprietary or financial interest in the outcome sign a statement, such as ‘I had full access to all of the data in this study and I take complete responsibility for the integrity of the data and the accuracy of the data analysis’. Editors may choose not to consider an article if a sponsor has asserted control over the authors’ right to publish (http://www.icmje.org/ethical_4conflicts.html).

5.5 The Effect of Bias on the Magnitude of Clinical Outcomes in Periodontology

It is clear that the reported methods of RCTs in periodontology are not optimal, and, therefore, the question arises, what is the effect of improper

trial methods on outcome size in these studies? Indirect evidence for an effect of bias in periodontology on outcome size was reported in a systematic review (Needleman et al. 2006) which found that when studies without both operator and examiner blinding were excluded from a meta-analyses, the difference between test and control became smaller and non-statistically significant. The small number of included trials suggests caution in drawing conclusions. However, when comparing the authors’ meta-analysis of GTR versus open flap debridement with previous meta-analyses which included studies at greater risk of bias, the greater the potential for bias in the previous meta-analyses, the greater the apparent benefit of the test treatment (difference in CAL gain was twice that comparing meta-analyses with most versus least risk of bias studies included). Fenwick et al. (2008) investigated the impact of allocation concealment and examiner masking on the size of clinical outcomes. Adequate allocation concealment and examiner masking were found in 24% and 64% of trials, respectively. There were no statistically significant differences in the magnitude of treatment outcomes comparing adequate versus inadequate or unclear allocation concealment (Table 5.3), nor in comparing adequate and inadequate examiner masked trials (Table 5.4). Retrospective power calculations suggest that at least 265 trials would be needed to demonstrate an effect if the magnitude of the overestimation of the magnitude of clinical effect was 0.5 mm. Future definitive research on

Table 5.3 Results of allocation concealment meta-analysis (Fenwick et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Effect of allocation concealment on PD	Effect size difference (mm) ^a	95% CI	SE	p value
Adequate versus unclear	0.22	−0.58, 1.03	0.41	0.59
Adequate versus inadequate	0.60	−1.70, 1.89	0.66	0.37
Adequate versus (either inadequate or unclear)	0.25	−0.53, 1.03	0.40	0.53
<i>Effect of allocation concealment on CAL/PAL</i>				
Adequate versus unclear	0.05	−0.95, 1.06	0.51	0.92
Adequate versus inadequate	−0.09	−2.0, 1.82	0.98	0.93
Adequate versus (either inadequate or unclear)	0.10	−1.01, 1.04	0.53	0.98

^aMean difference in treatment effect size according to study quality classification. Positive mean difference in treatment effect size indicates a tendency for poor quality studies to obtain greater treatment effect sizes. Conversely, a negative mean difference indicates good quality studies obtain greater treatment effect sizes

Table 5.4 Results of examiner masking meta-analysis (Fenwick et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Effect of allocation concealment on PD	Effect size difference (mm) ^a	95% CI	SE	p value
Adequate versus inadequate	-0.20	-0.76, 0.36	0.29	0.49
<i>Effect of examiner masking on CAL/PAL</i>				
Adequate versus inadequate	-0.19	-1.05, 0.68	0.44	0.67

^aMean difference in treatment effect size according to study quality classification

Table 5.5 Effect of outcome measure mean difference on study power and size (Fenwick et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Treatment effect size	2 mm mean difference between test and control				1 mm mean difference between test and control				0.5 mm mean difference between test and control			
Assessment criterion	Allocation concealment		Examiner masking		Allocation concealment		Examiner masking		Allocation concealment		Examiner masking	
Outcome measure	PD	CAL	PD	CAL	PD	CAL	PD	CAL	PD	CAL	PD	CAL
Effect size	0.93	0.68	1.32	0.82	0.46	0.34	0.66	0.41	0.23	0.17	0.33	0.20
% Observed power	99	95	99	99	67	45	92	59	22	15	39	19
No. of studies included in meta-regression	29	31	28	31	29	31	28	31	29	31	28	31
No. of studies required for 80% power of meta-regression	N/A	N/A	N/A	N/A	39	70	N/A	50	148	265	74	191

N/A not applicable as the current meta-regression already had over 80% with the existing number of studies for this effect size

this topic will therefore need to examine a much larger sample of trials (**Table 5.5**) (Fenwick et al. 2008).

References

- Demmer RT, Molitor JA, Jacobs DR Jr, Michalowicz BS. Periodontal disease, tooth loss and incident rheumatoid arthritis: results from the First National Health and Nutrition Examination Survey and its epidemiological follow-up study. *J Clin Periodontol*. 2011;38:998–1006
- DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, Kotzin S, Laine C, Marusic A, Overbeke AJ, Schroeder TV, Sox HC, Van Der Weyden MB, International Committee of Medical Journal Editors. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA*. 2004;292:1363–4.
- Dunn WR, Lyman S, Marx R; ISAKOS Scientific Committee. Research methodology. *Arthroscopy*. 2003;19(8):870–3.
- Faggion Jr CM, Listl S, Tu YK. Assessment of endpoints in studies on peri-implantitis treatment – a systematic review. *J Dent*. 2010;38:443–50.
- Fenwick J, Needleman IG, Moles DR. The effect of bias on the magnitude of clinical outcomes in periodontology: a pilot study. *J Clin Periodontol*. 2008;35:775–82.
- Flick SN. Managing attrition in clinical research. *Clin Psychol Rev*. 1988;8:499–515.
- Giannobile WV, Somerman MJ. Growth and amelogenin-like factors in periodontal wound healing. A systematic review. *Ann Periodontol*. 2003;8(1):193–204.
- Haines SJ. Randomized clinical trials in the evaluation of surgical innovation. *J Neurosurg*. 1979;51:5–11.
- Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615–25
- Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions* version 5.1.0. Updated March 2011. The Cochrane Collaboration. 2011. Available from www.cochrane-handbook.org.
- Hill EG, Slate EH, Wiegand RE, Grossi SG, Salinas CF. Study design for calibration of clinical examiners measuring eriodontal parameters. *J Periodontol*. 2006;77(7):1129–41.
- Hujoel PP, DeRouen TA. A survey of endpoint characteristics in periodontal clinical trials published 1988–1992, and implications for future studies. *J Clin Periodontol*. 1995;22(5):397–407.
- Hujoel PP. Endpoints in periodontal trials: the need for an evidence-based research approach. *Periodontol* 2000. 2004;36:196–204.
- Koch GG, Paquette DW. Design principles and statistical considerations in periodontal clinical trials. *Ann Periodontol*. 1997;2:42–63.
- Laine C, Horton R, DeAngelis CD, Drazen JM, Frizelle FA, Godlee F, Haug C, Hébert PC, Kotzin S, Marusic A,

- Sahni P, Schroeder TV, Sox HC, Van der Weyden MB, Verheugt FW. Clinical trial registration – looking back and moving ahead. *N Engl J Med*. 2007;356:2734–6.
- Leon AC, Mallinckrodt CH, Chuang-Stein C, Archibald DG, Archer GE, Chartier K. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *Biol Psychiatry*. 2006;59:1001–5.
- Lobo FS, Wagner S, Gross CR, Schommer JC. Addressing the issue of channeling bias in observational studies with propensity scores analysis. *Res Social Adm Pharm*. 2006;2:143–51.
- Loos BG, Tjoa S. Host-derived diagnostic markers for periodontitis: do they exist in gingival crevice fluid? *Periodontol* 2000. 2005;39:53–72.
- Lopez R, Scheutz F, Errboe M, Baelum V. Selection bias in case-control studies on periodontitis: a systematic review. *Eur J Oral Sci*. 2007;115:339–43.
- Lynch SE, Lavin PT, Genco RJ, Beasley WG, Wisner-Lynch LA. New composite endpoints to assess efficacy in periodontal therapy clinical trials. *J Periodontol*. 2006;77:1314–22.
- Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001a;285:1987–91.
- Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2001b;357:1191–4.
- Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med*. 2001c;134:657–62.
- Needleman IG, Worthington HV, Giedrys-Leeper E, Tucker RJ. Guided tissue regeneration for periodontal infra-bony defects. *Cochrane Database Syst Rev*. 2006;2:CD001724.
- Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. *Plast Reconstr Surg*. 2010;126:619–25.
- Paradis C. Bias in surgical research. *Ann Surg*. 2008;248(2):180–8.
- Pearce N, Checkoway H, Kriebel D. Bias in occupational epidemiology studies. *Occup Environ Med*. 2007;64:562–8.
- Petri H, Urquhart J. Channeling bias in the interpretation of drug effects. *Stat Med*. 1991;10:577–81.
- Polson AM. The research team, calibration, and quality assurance in clinical trials in periodontics. *Ann Periodontol*. 1997;2:75–82.
- Reynolds MA, Aichelmann-Reidy ME, Branch-Mays GL, Unsolley JC. The efficacy of bone replacement grafts in the treatment of periodontal osseous defects. A systematic review. *Ann Periodontol*. 2003;8(1):227–65.
- Schull WJ, Cobb S. The intrafamilial transmission of rheumatoid arthritis. III. The lack of support for a genetic hypothesis. *J Chronic Dis*. 1969;22:217–222.
- Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soc*. 2002;8:448–60.
- Storandt M, Morris JC. Ascertainment bias in the clinical diagnosis of Alzheimer disease. *Arch Neurol*. 2010;67:1364–9.

This chapter focuses primarily on the periodontal measurement sources of variability and on practical issues related to training, calibrating, and monitoring the individuals involved in conducting a clinical trial. It is also provided an overview of statistical methods used to perform evaluation of intra-examiner (comparisons may be made in which the same investigator examines the same subjects two or more times) and inter-examiner reliability (in which different investigators examine the same subjects). Details are also given regarding the reliability studies performed in periodontal research.

6.1 Basic Terminology

Accuracy refers to how close the results of a study are to the truth. The truth is often difficult to determine; furthermore, there may be controversy as to what constitutes the truth. An acceptable measure of truth to compare with radiographic bone height is measurements taken with a probe during periodontal surgery when the bone is exposed (Hausmann 2000).

Validity is a term sometimes used in periodontal literature, which is considered synonymous to accuracy (Hausmann 2000).

Reliability is defined as the degree to which multiple assessments of a subject agree (reproducibility) (Bartko 1991). Comparisons may be made in which the same investigator examines the same subjects two or more times (**intra-examiner reliability**) or in which different investigators

examine the same subjects (**inter-examiner reliability**) (Oakley and Brunette 2002).

6.2 Sources of Variability for the Measurement Reliability

Measurement reliability refers to the ability to obtain the same measurement consistency over sequential measures (Oakley and Brunette 2002). **Consistency** can be defined as the agreement of two quantitative measurements if it is assumed that neither of the two measurements is the gold standard. This is often encountered in calibration studies of periodontal disease where the observations of two or more examiners are compared, and true pocket depths or attachment levels are unknown (Hefti 1997).

The reliability of a measurement may be affected by three sources of variability: (1) the system or phenomenon being examined; (2) the examination itself, such as the instruments or equipment used and the examination environment; and (3) the examiners (Oakley and Brunette 2002).

6.2.1 Variation in the System or Phenomenon Being Measured

Normal biological variability may be inherent in the phenomenon being measured (Oakley and Brunette 2002). For example, it is well known that cortisol in serum varies diurnally with the highest

concentration in the morning (200–800 nmol/l) followed by a decrease during the day (<300 nmol/l) until midnight (Aardal and Holm 1995). Cortisol, like many other steroids, is also present in saliva and in gingival crevicular fluid (Axtelius et al. 1998). It was suggested that the total concentration of cortisol in gingival crevicular fluid might be estimated to levels below 1/10 of that in serum (Axtelius et al. 1998). Important determinants of salivary cortisol responses to stress in humans are age and gender, endogenous and exogenous sex steroid levels (e.g. the female menstrual cycle, use of oral contraceptives, and hormone replacement therapy), pregnancy, lactation and breastfeeding, smoking, coffee, and alcohol consumption as well as dietary energy supply in salivary cortisol responses to acute stress (Kudielka et al. 2009). In aim to reduce biological and methodological variation in salivary cortisol, it was recommended that (1) time of sampling has to be carefully registered and included in the statistical analysis; (2) samples have to be collected at the same time of year in longitudinal designs; (3) food intake has to be avoided in at least the 2 h before sampling; (4) vigorous exercise has to be avoided in at least the 2 h, preferably longer, before saliva is collected for measurement of cortisol; (5) variation in results obtained by different laboratory techniques emphasizes use of the same, or otherwise made comparable, laboratory techniques; (6) concentration of cortisol is dependent on the material of the tampon; (7) despite the absence of hard evidence, it is recommended that information be collected and results possibly statistically controlled for alcohol consumption, medication, such as oral contraceptives, and treatment for mental diseases; and (8) saliva samples can be stored at -20°C for at least 1 year (Hansen et al. 2008).

Interleukin- 1β (IL- 1β) is an important parameter in periodontal research because of its role in inflammation and bone resorption. One measure used to assess local IL- 1β concentrations is analysis of its levels in gingival crevicular fluid. While studies on serum IL- 1β concentrations indicate a circadian rhythm of this parameter, data showed a significant variation throughout the day of IL- 1β in GCF, with the lowest concentrations and

total amounts in the morning and the highest in the evening. The effect sizes of comparisons between morning and evening samples were medium to high and corresponded in magnitude to those reported in other published research comparing healthy sites and those affected by periodontitis. The smallest daytime variations were found to occur between 12:00 and 18:00 h. It is concluded that daytime variations in GCF-IL- 1β are large enough to be able to mimic or mask differences caused by clinical factors (Bergmann and Deinzer 2008).

In contrast, gingival crevicular fluid volume determination was revealed to be done with high reliability, the validity of measurements neither being affected by supragingival plaque nor by diurnal rhythms (Deinzer et al. 2000).

Analysis of intra- and inter-examiner reproducibility of probing depth measurements for a conventional manual probe (Williams) revealed that agreement for duplicate probing depth measurements of anterior teeth was significantly higher than posterior teeth ($P < 0.05$). Variability in proximal surfaces (especially distal surfaces) was significantly more than mid-facial or mid-lingual surfaces ($P < 0.05$). The difference between the agreements of mesial and distal surfaces was not significant ($P < 0.422$). Facial surfaces showed significantly more agreement than lingual surfaces ($P < 0.05$) (Lafzi et al. 2007). However, Fleiss et al. (1991) reported minor differences in variability between sites on anterior and sites on posterior teeth, between mid-sites and proximal sites, and between sites on the facial and sites on the lingual surface. When differences were found between the average variances for different kinds of sites, they may have been due to corresponding differences between the average depths and not to any inherently greater difficulty in measuring one kind of site than another.

6.2.2 Variability from the Instruments or Equipment Used and the Examination Environment

A number of probing methods and instruments have been developed in an attempt to address

limitations in obtaining clinical attachment-level measurement. First-generation instruments include conventional periodontal probes; second-generation probes utilize controlled forces; and third-generation probes incorporate automated measurement, controlled forces, and computerized data capture. Various types of stents have been used, and repeated measurement techniques have been proposed to reduce examiner error. Controlled-force probes appear to have their greatest advantage in increasing inter-examiner repeatability. The use of measurement stents increases inter- and intra-examiner reliability. However, use of such stents may be limited to small sample studies of limited duration. Third-generation instruments offer advantages in terms of automated measurement and data capture, increased resolution, and a more continuous measurement scale, but do not necessarily result in increased intra- or inter-examiner reliability. Decisions for or against use of a particular instrument must be made on the basis of the needs of each clinical trial (Pihlstrom 1992).

In laboratory-based measurements, instruments are typically calibrated against established standards, and the measurements are performed under controlled and specific conditions. The results and variability in these measurements are usually expressed as a standard deviation of the individual values or as confidence intervals around the calculated means (Oakley and Brunette 2002).

It is important to distinguish the reliability of a measurement from the precision of the measurement. The precision of the measurement refers to the exactness or degree of refinement with which a measurement is stated (Oakley and Brunette 2002).

The operation of measuring a distance, for example, reading a millimetre value from the markings of the periodontal probe, can be associated with a **digit preference error**. For example, the periodontal probe with Williams' markings allows for precise readings at 1, 2, 3, 5, 7, 8, 9, and 10 mm, but markings at 4 and 6 mm are missing. Examiners using this instrument tend to bias their readings towards an odd number, especially 5 mm, if the measured value falls anywhere between 3 and 7 mm. Such bias can be corrected

if the examiner is made aware of it or completely eliminated by the use of probes with markings at each millimetre or by electronic data capturing and digital read-out. In addition, the millimetre markings on periodontal probes are not always accurate (Hefti 1997). For example, clinicians may measure the depth of the periodontal pocket to the nearest millimetre with a periodontal probe. Although manufacturing methods to produce periodontal probes have been improved, a possible source of errors can reside in **the accuracy of their millimetre markings** (Winter 1979; Neto et al. 2001). **It is recommended that if you are uncertain how a probe is calibrated, use a millimetre ruler to determine the millimetre markings** (Nield-Gehrig 2007).

Van der Zee et al. (1991) investigated seven different probe types, representing currently marketed major designs (WHO-CPITN, Williams, Michigan), as well as available calibration systems (engraved, etched, and painted markings). Width of markings, accuracy of calibration from probe tip, and tine diameter at the tip and at specified points along the tine were assessed, using a stereomicroscope at a magnification of $\times 40$. Blind duplicate measurements of these probe tine characteristics were 100% reproducible to within 0.01 mm. There was an overall range in marking width from 0.00 to 1.13 mm. The best marking, in that it had no appreciable width and the highest accuracy, was the discrete transition between normal and engraved parts of probes with engraved bands. Mean inaccuracies of different probe sets varied from 0.06 to 0.22 mm. Probes from the same batch from the same production line could differ by more than 0.5 mm in calibration. Mean tip diameter ranged from 0.28 to 0.70 mm. The authors concluded that **probe tine diameter** and calibration should be considered in addition to other variables of periodontal probing. Standardization of tine characteristics and **avoidance of the use of different types or batches in a single study should enhance the accuracy and reproducibility of periodontal probe-dependent measurements** (Van der Zee et al. 1991).

Variability may also originate from the incorrect function or use of measuring devices or instruments. For example, reliable periodontal

Fig. 6.1 Clinical photos from the examination were a pressure-controlled TPS probe (TPS probe, Vivadent Ets., Schaan, Liechtenstein) with a diameter of 0.4 mm and calibrated for a 0.2N probing force was used. Probing depth registration at the buccal aspect of one experimental implant (a) and the right 1st mandibular molar (b) (Abrahamsson and Soldini 2006. Reprinted with permission from John Wiley & Sons, Inc.)

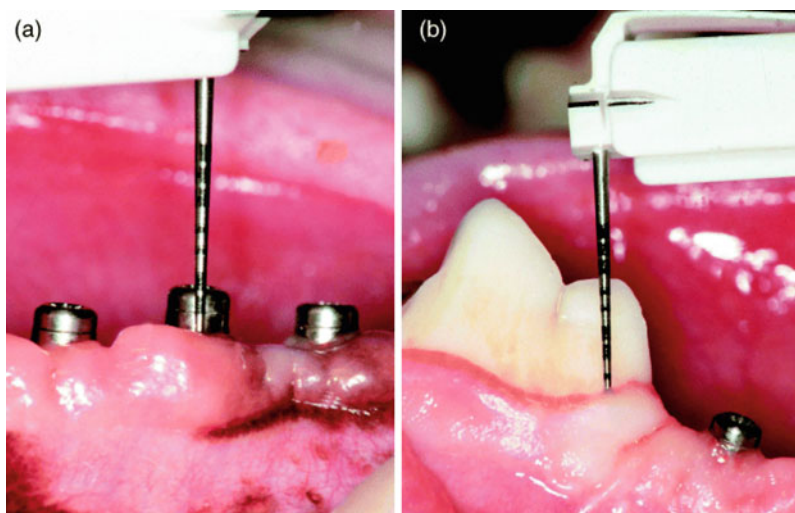
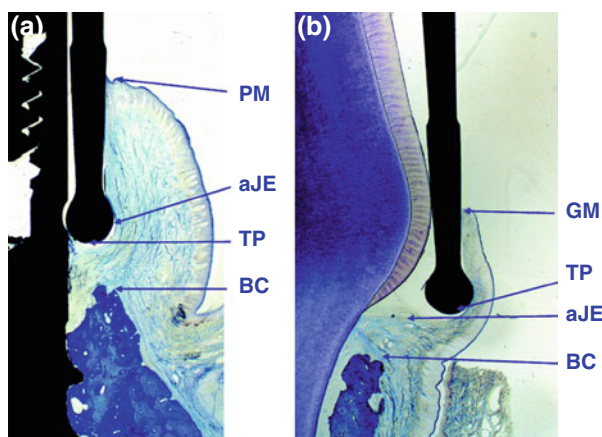


Fig. 6.2 Buccolingual cross-sections showing the metal probe tip position during peri-implant (a) and periodontal (b) probing (Abrahamsson and Soldini 2006. Reprinted with permission from John Wiley & Sons, Inc.)



probing requires the use of a calibrated probe, on correct positioning of the probe, and application of appropriate probing pressure (Oakley and Brunette 2002; Larsen et al. 2009). The majority of studies report only probing force while it is the pressure at the tip, the resultant of probing force and probe diameter, that eventually determines probe penetration (Figs. 6.1 and 6.2). Therefore, in a systematic review on the influence of probing pressure on the probing pocket measurements (PPD), an attempt was made to obtain a correction factor for comparing data obtained with different probing pressures. The probing pressures in the selected studies, all using tapered probe tines, ranged from extremely low to very high, that is, 51–995 N/cm². PPD increased with

increasing probing pressure in both diseased and healthy/treated sites. The correction factor compensating for the influence of the used probing pressure used at healthy/treated sites amounted to 0.002 mm per increase of 1 N/cm² in probing pressure, whereas at diseased sites, this was 0.004 mm (Table 6.1) (Larsen et al. 2009).

In order to overcome some of the variables associated with periodontal probing, several new probes have been invented: the Florida Probe System (Gibbs et al. 1988), the Toronto probe (Birek et al. 1987), Accutek probe (Goodson and Kondon 1988), Alabama probe (Jeffcoat et al. 1986; Jeffcoat 1994), etc. Inter- and intra-examiner variability using standard and constant force periodontal probes when measuring probing

Table 6.1 Summary of selected study divided into diseased and healthy/treated sites; mean probing pocket depth (PPD) per probing force/pressure (mm) and standard deviation in parenthesis (if available); increase in PPD (mm) calculation for each increase in probing force in relation to the preceding probing force (Larsen et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Author study	Forces (N)	Diameter (mm)	Pressure (N/cm ²)	PPD (diseased sites)	PPD increase	PPD (healthy/treated sites)	PPD increase
Bulthuis et al. (1998)	0.10	0.50	51	2.80 (1.88)	–	–	–
	0.15		76	2.83 (1.81)	0.03	–	–
	0.20		102	3.11 (2.00)	0.28	–	–
	0.25		127	3.14 (2.02)	0.03	–	–
Mombelli et al. (1997)	0.25	0.40	199	–	–	3.41 (0.49)	–
	0.50		398	–	–	3.92 (0.64)	0.51
	0.75		597	–	–	4.08 (0.72)	0.16
	1.00		796	–	–	4.16 (0.71)	0.08
	1.25		995	–	–	4.21 (0.69)	0.05
Barendregt et al. (1996)	0.25	0.50	127	–	–	2.40 (1.20)	–
	0.50		255	–	–	2.70 (1.40)	0.3
Chamberlain et al. (1985)	0.25	0.50	127	5.1 (1.4)	–	4.1 (1.3)	–
	0.50		255	6.1 (1.2)	1.0	4.7 (1.3)	0.6
	0.75		382	6.7 (1.0)	0.6	5.2 (1.3)	0.5
Caton et al. (1981)	0.15	0.35	155	3.06 (0.44)	–	2.00 (0.41)	–
	0.25		259	3.60 (0.51)	0.54	2.36 (0.41)	0.36
	0.50		520	3.99 (1.80)	0.39	2.64 (0.47)	0.28

depths was evaluated by Walsh and Saxby (1989). Thirty sites in ten patients with untreated chronic adult type periodontitis were examined by two operators, firstly using a standard periodontal pocket probe then with a constant force probe. Neither examiner was aware of the others readings and was thus blind in relation to one another. Overall, there was a maximum variation of ± 1 mm in 79.9% of recordings using the standard probe, and this agreement was increased to 100% with the use of the constant pressure probe. Similar results were reported by Magnusson et al. (1988) and Osborn et al. (1990). A clear advantage for both Florida probes is the capability of obtaining measurements to the nearest 0.1 mm, while with the conventional probe, measurements must be recorded to the nearest whole millimetre (Osborn et al. 1990; 1992).

Reddy et al. (1997) compared the intra-examiner and inter-examiner error of two constant force probes to the reading of a conventional manual probe. Three examiners made repeated examinations of attachment level using a modified Florida probe and a manual North Carolina probe (read to 1 or 0.5 mm); relative attachment-level

measurements were made using a Florida disk probe. Error was calculated as the mean of the absolute value of the difference between each examination, and the correlation between values at each examination was calculated. There was a significant difference in error by probe type (modified Florida probe 0.62 ± 0.03 mm, $r=0.86$; Florida stent probe 0.55 ± 0.05 mm, $r=0.82$; manual probe to 1 mm 0.39 ± 0.02 mm, $r=0.88$; manual probe to 0.5 mm 0.40 ± 0.02 mm, $r=0.89$; ($P<0.001$)). These data indicate that both manual and controlled-force probes can provide measurement within less than 1 mm of error; however, individual calibration of examiners remains important in the reduction of error.

6.2.3 Variability of Examiners

Examiners may be had varying levels of prior experience and little to no prior experience in clinical research. Examiners also differ because of biologic variation in the acuity of their senses (e.g. sight, touch, hearing), which may be further affected by their mood and sleep status.

A clinician's diagnosis may also be affected by the mind set; that is, clinicians tend to diagnose what they expect or hope to find (Oakley and Brunette 2002).

A study performed by Abbas et al. (1982) evaluated the effect of training upon inter-examiner agreement and concluded that only a training programme together with a standardized calibration of the probing force leads to reproducible probing depth measurements. Quirynen et al. (1993) observed that the intra-examiner deviation is smaller than the inter-examiner deviation and recommended that a **single examiner should follow the same subject during longitudinal clinical trials**.

Seabra et al. (2008) evaluated the influence of **examiner experience** on the variability of periodontal probing depth measurements obtained by conventional manual probing. 8,127 periodontal sites in 30 subjects with a diagnosis of chronic periodontitis were evaluated by an experienced examiner using an electronic probe and randomly assigned to three groups. Examiners with different levels of experience (undergraduate students, postgraduate students, and associate professors) evaluated each group with a manual probe. Electronic and conventional probing were repeated 45 days after cause-related periodontal therapy. The best agreement between electronic and manual probing at the baseline examination was obtained by the postgraduate students ($\kappa=0.66$) and at reassessment by the associate professors ($\kappa=0.60$). Undergraduate students obtained the lowest agreement values in both examinations ($\kappa=0.42$ and 0.11 , respectively).

6.2.3.1 Periodontal Examiner Training

Multiple examiners are often used in major national health surveys, large epidemiological studies, multi-centre studies, and various periodontal clinical trials. For periodontal disease, condition assessments of the principal outcome measures are usually performed with a standard mechanical probe. As differences in pressure, angulation, rounding, and other factors can produce different scores for the same site, it is essential that **all examiners be trained** in the

appropriate use of this probe using didactic sessions and physical application of the procedures. A **diagnostic manual** completely describing and illustrating the exact procedures for disease diagnosis should be prepared and made available to each examiner (Kingman and Albandar 2002). This standard operating procedure document will serve as a reference source during the didactic, clinical training, and actual clinical study components. For example, when using the periodontal probe for probing depth determination, specific details should be given regarding technique, alignment, location points for measurements, and measurement rules when the gingival margin falls between graduation of the probe. Furthermore, attachment-level determinants incorporate an additional component, namely, a location of the gingival margin to a fixed reference point, for example, the cemento-enamel junction (CEJ), and there needs to be a separate consideration of relevant component such as anatomic awareness, tactile discrimination, and management of situations when the CEJ is either subgingival or supragingival (Polson 1997). Prior to the conduct of the survey, adequate training and evaluation of the examiners should be conducted to document that the examiners are scoring diseases accurately and consistently. The use of an expert reference examiner is an integral component of the quality control process (Kingman and Albandar 2002).

Although there is near universal consensus within the oral health research community as to the importance of examiner reliability, the literature is sparse with respect to appropriate design of examiner calibration studies (Hill et al. 2006).

For periodontal clinical trials, Polson (1997) suggests using data from five subjects to assess intra-examiner variability and data from between three and five subjects to assess reliability between each examiner and a designated standard examiner. The first procedures focus on technique considerations, rather than measurement reproducibility considerations. The latter should only be tested after all of the technique-related components have been satisfactorily achieved. The repeatability of measurement testing follows as a separate component after satisfactory completion

of the technique components. Examiners should make these measurements in liaison with a chair-side recorder so communication skills can be developed between team members. Repeat measures should be done on several different participants, using only one quadrant from each. In this way, the examiner gains initial experience within and between subjects in clinical variation. Repeat measure data on a particular subject should not be obtained until at least 15 min have elapsed since completion of the first exam, thereby minimizing examiner memory recollection of previous recordings. The data set which is obtained from this initial calibration exercise can be analysed to give trends for guidance; however, it should not be used as an example of a database from controlled conditions. The formal database for evaluating intra-examiner variability should be obtained as a controlled clinical exercise according to details from standard operating procedures (Polson 1997).

The examiner usually aggregates this database over a specific number of subjects within a specified window of time after the core instructor has left the study centre. It is customary to aggregate these data from five different subjects over a 2- to 3-week time period. These completed data sets are analysed for measures of variability. A decision is then made regarding the acceptability of this variability in relation to the levels defined by the sponsor and the standards for clinical trials from the periodontal literature. Assuming that acceptable levels have been met, the final stage in the training and calibration of the periodontal examiner is achieved by the core instructor revisiting the research centre facility and fine-tuning examiner methods and procedures. This joint session incorporates review of the database of repeated measures, reviews of clinical technique by observing the examiner collecting data on subjects, and an inter-examiner comparison between the research centre examiner and the core instructor by comparing measurements from the same subjects (usually 3–5). Satisfactory completion of these components qualifies the periodontal examiner to participate in the next stages of the team and centre framing programme (Polson 1997).

In the NHANES III survey, detailed multilevel quality control procedures were built into the protocol. A reference examiner was hired to train the survey examiners and monitor them throughout the survey period. The examiners were evaluated before the survey began and were also monitored during the survey period. Two types of examiner evaluations involving replicate examinations of subsets of study participants are available for the NHANES III. The first type of evaluation consisted of intra-examiner reliability. The intra-examiner assessment was based on roughly 20 sampled persons at each survey location who were scheduled for re-examination by the same examiner in the medical examination centre. The replicate examinations were conducted within a 3- or 4-week period, and the cumulative replicate data comprise the intra-examiner reliability database used in the analyses. The second type of evaluation procedure involved periodic visits to examination sites by the reference examiner. During these visits, the expert examiner conducted replicate periodontal examinations on a subset of survey participants. These examinations were generally conducted on the same day as those by the survey examiner. The cumulative set of replicate examinations obtained during these visits constituted the inter-examiner reliability database used in the analyses. The total number of subjects across all locations used to assess examiner reliability ranged from 76 to 304 (intra) and from 48 to 153 (inter) for the three examiners whose examinations accounted for 94% of the database (Kingman and Albandar 2002).

Recently, Hill et al. (2006) presented an approach to examiner calibration study design where the number of calibration subjects is based on a specified margin of error (half-width of the 95% confidence interval [CI]) of the percentage of agreement (exact and within 1 mm) for both intra- and inter-examiner reliability assessments. An experienced standard examiner (S) trained three dental hygienists (A, B, and C) in correct procedures for obtaining a variety of periodontal measures. Duplicate measurements of probing depth (PD [mm]) and the free gingival margin to the cemento-enamel junction (CEJ-GM [mm])

were obtained in a pilot study to design a formal examiner calibration study, where sample sizes were adjusted for the effects of within-subject clustering of binary indices of agreement. Specifically, the authors designed their calibration study based on estimation of percentage of agreement and agreement within 1 mm. Within-subject clustering of agreement indices resulted in an approximate fourfold increase in the variance of the estimates of percentage of agreement with the standard. PD and CEJ-GM percentage of exact agreement measurements (95% CI) for each examiner-standard pair, respectively, were as follows: AS = 55% (48%, 61%) and 70% (62%, 78%); BS = 52% (45%, 59%) and 73% (63%, 82%); and CS = 55% (50%, 61%) and 72% (65%, 79%). The corresponding 95% CIs unadjusted for the effects of clustering underestimated the margin of error associated with the estimates of exact agreement by as much as 57% for PD and 68% for CEJ-GM (Hill et al. 2006).

6.2.3.2 Therapist Training

As part of the team and centre training, the therapist who is administering an experimental treatment may need a special training programme. This programme, like the programme for the periodontal examiner, should have a didactic, laboratory bench-top, and clinical administration portion, as well as an evaluation of clinical technique components. The format for the programme can be designed in a similar manner to that for the examiner training programme (Polson 1997).

6.2.3.3 Team Training Protocol

After the training programmes have been satisfactory completed for the periodontal examiner and the therapists for the experimental treatment and the conventional treatment, the entire research team should complete a training protocol which is a mini-facsimile of the protocol for the actual multi-centre investigation. In this training protocol, the study centre should enrol four to six subjects so the research team gains collectively experience in patient enrolments, study procedures, case-report form managements, and regulatory requirements. The training protocol is monitored in the same manner than the eventual

clinical trial will be monitored. Each study centre team needs to satisfactory complete this training protocol before being authorized to participate in the actual multi-centre study (Polson 1997).

6.2.3.4 Study Execution

Monitoring procedures should be established so that the study investigators know and understand what is expected of them during the study. Periodic review of the important components of the study procedure should be an integral part of the study protocol. This includes monitoring the examiners' ability to consistently diagnose the clinical conditions under study (Kingman and Albandar 2002).

As investigators may need to make judgements using criteria that are not very specific or make judgements about subject characteristics that are difficult to evaluate, the best that can be done is to determine if the investigators are consistent in their judgements. That is, performance review of the clinician investigators focuses on the likelihood that repeated examinations of the same, unchanged patient by either the same clinician or other clinicians yield identical results (Oakley and Brunette 2002).

6.3 Standardization Versus Calibration

Standardization of clinical examiners should be performed prior to a planned study in order to minimize the impact of examiner style. For multi-centre studies where multiple examiners are employed, such standardization is of utmost importance and should include not only the evaluation of the various index systems or measurements used but also a detailed understanding of the procedures to be performed during the planned study. Calibration of an examiner follows standardization exercises. In this process, the reproducibility, rather than the accuracy of an examiner to detect signs of health and/or disease, is evaluated. Calibration without prior standardization may falsify the image of the true conditions present. The examiner may reach high reproducibility with an index system without correctly appraising

the relevant clinical changes (Lang et al. 2010; Egelberg 1999).

Intra-examiner calibration is influenced by examiner style. The ability to simply repeat scores will be higher when an examiner possesses a style that predominantly uses a pair of scores (style 0, 1; style 0, 2; or style 1, 2). In contrast, the use of kappa statistics to judge intra-examiner agreement on repeat measurements will tend to result in a higher kappa statistics when more of the index is used. Thus, an awareness of the examiner's style is necessary to more adequately interpret the results of a calibration exercise (McClanahan et al. 2001).

The studies presented by McClanahan et al. (2001) were performed in aim to develop an understanding of how clinicians experienced with Löe–Silness gingival index (GI) differ with respect to how they apply GI and to assess the impact of different examination styles on statistical outcomes and magnitude of treatment differences. Retrospective analyses indicated the presence of distinct examiner styles which are based on the frequency that a given GI score (0, 1, 2, or 3) is measured by a clinician. In the prospective study, all 12 examiners observed statistically significant differences between the prophylaxis treatment groups at the final visit for both mean number of bleeding sites and mean GI; the magnitude ranged from 21.5% to 84.6% for mean number of bleeding sites and 9.4–39.2% for mean GI. There were four distinct styles employed by these experienced clinicians.

These data support the use of examiner calibration as a mechanism to standardize clinical examiners and minimize the impact of examiner style within the context of single-centre or multi-centre studies where multiple examiners are employed. However, there are significant obstacles to effectively dealing with this issue. There is no established objective standard to which examiners can be calibrated for GI because of the subjective nature of the index, and as such, typically a known standard examiner is arbitrarily defined as the 'gold standard'. This selection is routinely based on the examiner's ability to detect known treatment differences between products or their reproducibility as opposed to their accuracy

in detecting clinical inflammation. When examiner calibration is successfully completed, a common style can often be achieved. However, the examiner style which is most appropriate is not readily apparent. In the prospective mentioned above, differences in examiner style did not preclude the examiners from observing statistically significant treatment differences (prophylaxis effect) for mean number of bleeding sites and mean GI. While examiners employed the GI with dramatically different styles, their results consistently supported the same conclusion with respect to the statistical significance of the treatment effect (McClanahan et al. 2001).

While standardization of new examiners to experienced examiners generally requires large patient cohorts, the assessment of intra-examiner agreement may be performed with small number of subjects chosen appropriately for the purpose of the subsequent study. Often, repeated assessments are made during the actual study with a limited number of patients. While not ideal from an ideal and standardization point of view, this method of duplication during the study may provide adequate data to determine intra-examiner reproducibility of an experienced examiner (Lang et al. 2010).

6.4 Reliability Measures

Reliability is defined as the degree to which multiple assessments of a subject agree (reproducibility) (Bartko 1991). Comparisons may be made in which the same investigator examines the same subjects two or more times (**intra-examiner reliability**) or in which different investigators examine the same subjects (**inter-examiner reliability**). Inter-observer variability is minimized when the endpoints are well defined and quantifiable, such as measuring the anatomic root length on a periapical radiograph. Inter-observer variability is grater when criteria are vague and subjective (Oakley and Brunette 2002), such as visual–tactile indices, exclusively based on clinical presentation of gingiva and include visual assessment of oedema, colour, tissue form and contour, and/or an assessment of bleeding upon provocation (McClanahan et al. 2001). Bleeding following

provocation is often considered to be more objective and clinically relevant than assessments of gingival tissue appearance (colour and form) (McClanahan et al. 2001).

Reliability measures provide information only about how well the examiners agree, not about whether the conclusions are correct. Inter-examiner reliability and intra-examiner reliability have been quantitated by different measures: percentage agreement, correlation coefficients, intra-class correlation coefficient, and kappa values (Oakley and Brunette 2002).

There is increasing awareness among researchers that the two most appropriate measures of reliability are the intra-class correlation coefficient (for continuous variables) and kappa (for categorical variables). Fleiss and Gordon (1973) and Rae (1988) established correspondences between various kappa statistics and intra-class correlation coefficients under general conditions (multiple raters and polychotomous category systems).

However, unacceptable statistical measures of reliability such as chi-square, per cent agreement, product moment correlation, as well as any measure of association and Yule's Y still appear in the literature (Bartko 1991).

When evaluating the statistical results, Bartko (1991) also revealed that it is not acceptable, for example, to report simply that 'the reliability was 0.7'; it is essential to state the name of the reliability measure and perhaps even how it was computed. If it is an inappropriate measure, the reader will be in a position to make a judgement regarding quality. Results of studies are difficult enough to interpret without the added burden of unsuitable statistics. There are costs associated with improper measurements, unreliable diagnostic systems, inappropriate statistics and measures of reliability, and poor quality research. Costs are incurred when misleading information directs resources and talents into non-productive avenues of research (Bartko 1991).

6.4.1 Percentage Agreement

Several studies reporting levels of examiner agreement are using statistical measures, such as percentage agreement, which do not take

into account the possibility that some agreement occurs as a result of chance alone (Blieden et al. 1992; Walsh and Saxby 1989; Suh et al. 2002). When examiners are clinicians making diagnoses of disease versus health, as the proportion of true negative diagnostic decisions increases, the per cent of inter-examiner agreement due to chance also rises. This effect, for example, inflates estimates of agreement in caries surveys where a 5% or lower prevalence of dental caries assures a high level of overall per cent agreement, even though agreement on the presence of caries may be very poor (Valachovic et al. 1986).

For binary measures (e.g. plaque index), percentage of agreement is the most intuitive and interpretable to a broad audience but fails to adjust for chance agreement. For ordinal variables (e.g. calculus index with levels of none, supragingival, and either subgingival only or both subgingival and supragingival; or probing depth/attachment loss in 1-mm increments), percentage of agreement remains an option but treats all non-agreements equivalently. For example, differences in probing depths of 1 and 5 mm are equally incorrect. Alternatives include reporting agreement within 1 mm (for integer values of PD or attachment loss) or weighted kappa (κ_w) (Hill et al. 2006).

6.4.2 Standard Error of Measurement

Measurement error is estimated using the differences between the two values obtained for the same periodontal condition by an examiner (Kingman and Albandar 2002). The generic formula used for estimating measurement error at an individual site is derived from the model:

$$y_{ijr} = y_{ij} + \epsilon_{ijr}$$

where y_{ijr} is the observed value, y_{ij} is the true value, and ϵ_{ijr} is the measurement error. Here $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$; and $r = 1, 2$, with i indexing the subject, j the site within subject, and r the examination sequence for the subject. At the site level, one assumes that $E(\epsilon_{ijr}) = 0$ and $\text{var}(\epsilon_{ijr}) = \sigma_{ij}^2$. Each σ_{ij} is estimable, and its estimator is based on

the difference between replicate measurements d_{ij} , where

$$d_{ij} = y_{ij1} - y_{ij2}$$

If one assumes measurement errors are independent of subject and site type, that is, $\text{var}(y_{ijk}) = \sigma^2$, one can obtain an overall measurement error estimate by computing

$$s = \sqrt{\frac{1}{2kn} \sum (y_{ij1} - y_{ij2})^2} = \sqrt{\frac{1}{2kn} \sum d_{ij}^2}$$

(Kingman & Albandar, 2002).

6.4.3 Correlation Coefficients

The correlation coefficient is used with continuous data, and it is based on the extent to which the relationship between two variables can be described by a straight line called the regression line. The Pearson correlation coefficient, r , is a measure of the strength of the relationship between two sets of data. The strongest positive correlation has a value of 1.0, no relationship is indicated by 0, and perfect negative correlation has a value of -1.0 . Thus, correlation coefficients with values closest to 1.0 demonstrate the greatest relationship between sets of data, but perfect agreement occurs only when the regression line has a slope of 1; that is, the points fall along the line of equality (Oakley and Brunette 2002).

In regression analysis, the square of the correlation coefficient, r^2 , is, in effect, the fraction or proportion of the total variance in the dependent variable that can be explained by the relationship between the variables; r^2 tends to overestimate the true reliability. In general, r values will be greater than, and overestimated in comparison with, the more theoretical sum reliability calculated by the intra-class correlation coefficient (ICC) (Oakley and Brunette 2002).

Correlations (Pearson, Spearman, etc.) are inappropriate, as they measure association, not agreement. The Pearson correlation is unity for data set 1 (as well as the other two data sets) only because the rater data are linearly associated. The peculiar nature of the data, that is, that the raters agree perfectly (zero within-subject variance), has no effect on the

Pearson correlation. Chi-square measures also association, not agreement (Bartko 1991).

6.4.4 Intra-class Correlation Coefficient (ICC)

The intra-class correlation coefficient (ICC) of reliability measures the degree of correlation among observations within a class. In this context, the class is the collection of examiner pair responses. As a reliability measure, the ICC is interpretable as the proportion of total variation attributable to between-subject variability (Hill et al. 2006).

The ICC is generally derived from analysis of variance calculations. There are numerous versions of the intra-class correlation coefficient (ICC) that can give quite different results when applied to the same data. The guidelines for choosing the appropriate form of the ICC call for three decisions: (a) Is a one-way or two-way analysis of variance (ANOVA) appropriate for the analysis of the reliability study? (b) Are differences between the judges' mean ratings relevant to the reliability of interest? (c) Is the unit of analysis an individual rating or the mean of several ratings? The first and second decisions pertain to the appropriate statistical model for the reliability study, and the second and the third pertain to the potential use of its results (Shrout and Fleiss 1979). Shrout and Fleiss (1979) discussed six forms of the intra-class correlation and guidelines for choosing among them. Important issues in the choice of an appropriate index include whether the ANOVA design should be one way or two way, whether raters are considered fixed or random effects, and whether the unit of analysis is a single rater or the mean of several raters (Shrout and Fleiss 1979).

Intra-class correlation coefficient values can range from 0 to 1.0. Unlike r , the ICC value indicates what proportion of the total observed variability is caused by variability among the subjects as compared with variability among the examiners. If most of the variability results from discrepancies among examiners, the ICC values are low. Alternatively, if the examiners are reliable (consistent) among themselves, ICC values are high (e.g. between 0.75 and 1.00), and in effect, one examiner could be replaced with another. The

ICC values may be interpreted in a manner similar to κ scores, which are more commonly used (Oakley and Brunette 2002).

ICC can be expressed according to Goulet and Clark (1990) as:

- ‘Unacceptable’ (values between 0.00 and 0.39)
- ‘Moderate’ (values between 0.40 and 0.59)
- ‘Acceptable’ (values between 0.60 and 0.79)
- ‘Excellent’ (values between 0.80 and 1.00)

The intra-class correlation coefficients were used to measure the inter- and intra-examiner reliability for measurements of probing depth and attachment level (Albandar and Kingman 1999; Albandar et al. 1999; Preshaw et al. 1999; Rise and Tollefsen 1984; Marks et al. 1993), alveolar bone levels on radiographs (Hou et al. 2003; Teeuw et al. 2009; Pecoraro et al. 2005; Pierro et al. 2008; Sakka et al. 2005; Gröndahl et al. 1998; Müller and Ulbrich 2005), furcations (Eickholz et al. 1999), and oral hygiene measurements (Albandar and Kingman 1999; Rise and Tollefsen 1984; Marks et al. 1993).

6.4.5 Kappa Scores

The statistics kappa (Cohen 1960) and weighted kappa (Cohen 1968) were introduced to provide coefficients of agreement between two raters for qualitative or dichotomous judgements. Kappa is appropriate when all disagreements may be considered equally serious, and weighted kappa is appropriate when the relative seriousness of the different possible disagreements can be specified (Fleiss et al. 1969).

6.4.5.1 Cohen’s Kappa

Cohen (1960) suggested that, for any problem in nominal scale agreement between two judges, there are only two relevant quantities:

p_o = the proportion of units in which the judges agreed.

p_c = the proportion of units for which agreement is expected by chance.

The test of agreement comes then with regard to the $1 - p_c$ of the units for which the hypothesis of no association would predict disagreement between the judges. This term will serve as the denominator (Cohen 1960).

To the extent to which non-chance factors are operating in the direction of agreement, p_o will exceed p_c ; their difference, $p_o - p_c$, represents the proportion of the cases in which beyond-chance agreement occurred and is the numerator of the coefficient (Cohen 1960).

The coefficient κ is simply the proportion of chance-expected disagreements which do not occur, or alternatively, it is the proportion of agreement after chance agreement is removed from consideration:

$$\kappa = \frac{p_o - p_c}{1 - p_c}.$$

When obtained agreement equals chance agreement, $\kappa=0$. Greater than chance agreement leads to positive values of κ ; and less than chance agreement leads to negative values. The upper limit of κ is +1.00, occurring when (and only when) there is perfect agreement between the judges (Cohen 1960).

Cohen (1960) also presented large sample formulae for the standard error of an observed κ :

$$\sigma_{\kappa} = \sqrt{\frac{p_o(1 - p_o)}{N(1 - p_c)^2}}$$

used for setting confidence limits and performing two-sample hypothesis tests, and the standard error of κ when the population $\kappa=0$

$$\sigma_{\kappa_o} = \sqrt{\frac{p_c}{N(1 - p_c)}}$$

used for one-sample significance tests of κ . For large samples, the sampling distribution of κ is approximately normal, and statistical tests using normal curve deviates take the familiar classical form. A rule of thumb is that clinical studies should not proceed before investigators have been trained and calibrated with demonstrated high κ scores (e.g. $\kappa>0.6$).

The further development to κ_w (**weighted kappa**) is motivated by studies in which it is the sense of the investigator that some disagreements in assignments, that is, some off diagonal cells in the $\kappa \times \kappa$ matrix, are of greater gravity than others. The κ makes no such distinction, implicitly treating all disagreement cells equally. The κ_w represents the proportion of *weighted* agreement

corrected for chance to be used when different kinds of disagreement are to be differentially weighted in the agreement index (Cohen 1968).

The use of Cohen's kappa and weighted kappa is restricted to the case both where the number of raters is two and where the same two raters rate each subject. Generalizations are therefore necessary for the case of more than two raters and for the case where the raters judging one subject are not necessarily the same as those judging another (Fleiss 1971; Fleiss and Chilton 1983).

Fleiss (1971) extended kappa to the case in which each of a sample of subjects is rated on a nominal scale by the same number of raters but in which the raters rating one subject are not necessarily the same as those rating another. Whereas Cohen's kappa works for only two raters, **Fleiss' kappa** works for any number of raters giving categorical ratings to a fixed number of items. It can be interpreted as expressing the extent to which the observed amount of agreement among raters exceeds what would be expected if all raters made their ratings completely randomly. It is important to note that whereas Cohen's kappa assumes the same two raters have rated a set of items, Fleiss' kappa specifically assumes that although there are a fixed number of raters (e.g. three), different items are rated by different individuals. That is, item 1 is rated by raters A, B, and C, but item 2 could be rated by raters D, E, and F (http://en.wikipedia.org/wiki/Fleiss%27_kappa).

Let N represent the total number of subjects, n the number of ratings per subject, and k the number of categories into which assignments are made. Let the subscript i , where $i=1, \dots, N$, represent the subjects, and the subscript j , where $j=1, \dots, k$, represent the categories of the scale (Fleiss et al. 1979).

Define n_{ij} to be the number of raters who assigned the i th subject to the j th category and define

$$p_j = \frac{1}{Nn} \sum_{i=1}^N n_{ij}.$$

The quantity p_j is the proportion of all assignments that were to the j th category. The measure of the extent of agreement beyond chance in assigning subjects to category j ($j=1, \dots, k$) is

$$\kappa_j = 1 - \frac{\sum_i n_{ij}(n - n_{ij})}{Nn(n-1)p_j q_j}$$

where $q_j = 1 - p_j$. If there is perfect agreement in the assignments to category j (i.e. if each $n_{ij}=0$ or n), then $\kappa_j = 1$. If, on the other hand, the n_{ij} 's vary as binomial random variables with parameters n and p_j , then the expected value of κ_j is 0. The minimum value of κ_j is $-1/(n-1)$ (Fleiss et al. 1979).

The overall measure of agreement beyond chance is a weighted average of the κ_j s:

$$\kappa = \frac{\sum p_j q_j \kappa_j}{\sum p_j q_j} = 1 - \frac{Nn^2 - \sum \sum n_{ij}^2}{Nn(n-1) \sum p_j q_j}.$$

The overall measure also varies from a minimum of $-1/(n-1)$ for poorer than chance agreement through 0 for just chance agreement to unity for perfect agreement (Fleiss et al. 1979).

The general form of the test for comparing two independent kappa values is given by the critical ratio:

$$z = \frac{\kappa_1 - \kappa_2}{\sqrt{V_{(\kappa_1)} + V_{(\kappa_2)}}}$$

in which

- κ refers to a given kappa (or weighted kappa) value.
- V refers to the variance of a given kappa (or weighted kappa) value.
- Z is the value used to determine P levels from tables of the standard normal distribution (e.g. a Z value of 1.96 corresponds to a P level of 0.05, based upon a two-tailed test of statistical significance) (Cicchetti and Heavens 1981).

What does a specific kappa value mean? According to Viera and Garrett (2005), it is necessary to remember that perfect agreement would equate to a kappa of 1, and chance agreement would equate to 0:

Kappa value <0 : less than chance agreement
 Kappa value 0.01–0.20: slight agreement
 Kappa value 0.21–0.40: fair agreement
 Kappa value 0.41–0.60: moderate agreement
 Kappa value 0.61–0.80: substantial agreement
 Kappa value 0.81–0.99: almost perfect agreement

When interpreting kappa, it is also important to keep in mind that the estimated kappa itself could be due to chance. To report a P value of a

kappa requires calculation of the variance of kappa and deriving a z statistic. A confidence interval for kappa, which may be even more informative, can also be calculated. Fortunately, computer programs are able to calculate kappa as well as the P value or confidence interval of kappa at the stroke of a few keys. Remember, though, the P value in this case tests whether the estimated kappa is not due to chance. It does not test the strength of agreement. Also, P values and confidence intervals are sensitive to sample size, and with a large enough sample size, any kappa above 0 will become statistically significant (Viera and Garrett 2005).

The kappa statistics have been used by various studies to measure the inter- and intra-examiner reliability in periodontal research (Mariath et al. 2008; Albandar and Kingman 1999; Albandar et al. 1999; Kida et al. 2006; Mojon et al. 1996; Verhoeven et al. 2000; Spolsky and Gornbein 1996; Blieden et al. 1992; Eaton et al. 1997; Shi et al. 1999; Deas et al. 2006; Eickholz et al. 1999; Kawamura et al. 2000; Rise and Tollefsen 1984; Dombret et al. 2003; Wolf et al. 2001).

Spolsky and Gornbein (1996) compared methods used to assess reliability for gingival inflammation and plaque. Duplicate examinations were conducted by one examiner on 17 subjects (506 scoring sites), using the gingival index (GI), bleeding points index (BPI), and plaque index (PI). The percentage of agreement, the weighted and unweighted kappa coefficients, and Pearson correlation coefficients were calculated as statistics of reliability for mesial buccal site scores and whole mouth mean scores when appropriate. For mesial buccal sites, the respective values of the GI, BPI, and PI for weighted kappas were 0.47, 0.49, and 0.75; for the correlation coefficients 0.47, 0.49, and 0.76; for unweighted kappas 0.39, 0.49, and 0.39; and for percentage of agreement 66.2%, 76.1%, and 51.2%. For whole mouth means, the correlation coefficients for the GI, BPI, and PI were 0.87, 0.59, and 0.87, respectively. The authors concluded that the most useful statistics in assessing the intra-examiner

reliability of a solo examiner in descending order were the weighted kappa coefficient, Pearson correlation coefficient, the unweighted kappa coefficient, and percentage of agreement. The authors could demonstrate, indeed, that the percentage of agreement is sensitive to the number of ranks used to measure (i.e. the chance of agreement decreases as the number of measurement categories increases), while κ statistics and correlation coefficients do not. Moreover, percentage of agreement, unlike κ and Pearson correlation coefficients, is not corrected or controlled for chance agreement (Spolsky and Gornbein 1996). The Pearson correlation coefficient should, however, be preferred to the unweighted κ coefficient as it gives credit to 'near misses' disagreements (i.e. disagreements that are close to each other). Spolsky and Gornbein (1996) have also shown that when the k coefficient formula is corrected to give a higher value to a near miss, the calculated κ correlation coefficient tends to come closer to the Pearson correlation coefficient and that the more correction is given to near misses, the closer the weighted κ correlation coefficient comes to the Pearson correlation coefficient, resulting even mathematically almost equivalent (Dombret et al. 2003).

6.5 Inter- Versus Intra-examiner Variability

In multi-centre studies, multiple examiners are usually involved in assessing the same parameters. It is, therefore, a prerequisite to determine inter-examiner variability following standardization exercises, if results of the study are to be collapsed into one cohort, thereby increasing the statistical power of the study. Usually, calibration focuses on patient-based mean scores rather than single-site scores (Lang et al. 2010). Clinical studies on accuracy and reproducibility of periodontal probes have revealed that inter-examiner agreement is generally lower than intra-examiner agreement both for non-para-

metric indices as well as for continuous parameters (Lang et al. 2010; Osborn et al. 1990; Grossi et al. 1996; Mayfield et al. 1996; Khocht and Chang 1998). This fact reinforces the conviction that either a single examiner should do the follow-up probings particularly in longitudinal studies, or different examiners should be very well calibrated before taking part in the measurements (Buduneli et al. 2004).

Similarly, in experimental animal models, for example, the critical size, supra-alveolar, and periodontal defect model, a number of histologic outcomes are examined following treatments. Many of these outcomes are continuous-scale variables, and the more important of these include defect height and area and newly formed alveolar bone, height, and area (**Fig. 6.3**). Histometric analysis of the latter variables is performed regularly in studies of novel periodontal and implant therapies (Koo et al. 2004). Koo et al. (2004) showed that, when a single examiner performed the measurement of these parameters, between 1% and 6% of the total variation was due to examiner random measurement error. Furthermore, a measurement error component of about 1–17% of the total variation was due to differences between examiners (**Table 6.2**). This suggests that assessments made by one examiner are quite adequate for the analysis of data. The use of more than one examiner to perform the histometric examination may introduce an additional, and perhaps larger, error component.



Fig. 6.3 Examples of histometric parameters evaluated in the critical size, supraalveolar periodontal defect model. The light green line represents the base of the defect and the orange arrowheads, the cemento-enamel junction. The defect height (dark green arrow), bone regeneration height (yellow arrow), membrane height (purple arrow), defect area (light blue lines), and bone regeneration area (orange lines) are shown (Koo et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

for patient-based mean scores is higher than that for site-based single scores. This is, again, the case for non-parametric indices as well as for continuous measurements (Lang et al. 2010).

6.6 Intra-examiner Variability in Periodontal Studies

Determination of intra-examiner agreement has to follow standardization exercises prior to any clinical trial. Intra-examiner agreement procedures may be performed with patient-based scores or with site-based analysis. Both aspects may be of importance depending on the study planned. Obviously, intra-examiner agreement

6.6.1 Non-parametric Indices

These include most of the parameters recommended to evaluate oral hygiene levels, gingival health status, plaque retentive factors, and assessment of stain. Since these parameters are categorical, they may be best analysed applying frequency analyses instead of only mean scores.

Table 6.2 The intra-class correlation coefficients of repeated measurements of histometric parameters made by one examiner (intra-examiner reproducibility) or two examiners (inter-examiner reproducibility) (Koo et al. 2004) (Reprinted with permission from John Wiley & Sons, Inc.)

Parameter ^a	Intra-examiner	Inter-examiner
Defect height	0.985	0.908
Defect area	0.991	0.986
Membrane height	0.981	0.963
Junctional epithelium	0.976	0.761
Connective tissue repair	0.975	0.844
Bone regeneration (height)	0.882	0.834
Bone regeneration (area)	0.941	0.926
Bone regeneration (density)	0.771	0.731
Biomaterial density	0.661	0.605
Root resorption	0.994	0.892
Ankylosis	0.983	0.766

^a*Defect height*: distance between the apical extension of root planing and the cemento-enamel junction; *Defect area*: area under the ePTFE membrane circumscribed by the planed root, the width of the alveolar bone at the apical extension of the root planing, and the membrane; *Membrane height*: distance between the apical extension of the root planing and the most coronal aspect of the ePTFE membrane; *Junctional epithelium*: distance between the apical and coronal aspect of a junctional epithelium along the planed root; *Connective tissue repair*: distance between the apical extension of the root planing and the apical extension of the junctional epithelium along the planed root; *Cementum regeneration*: distance between the apical extension of the root planing and the coronal extension of a continuous layer of new cementum or a cementum-like deposit on the planed root; *Bone regeneration (height)*: distance between the apical extension of the root planing and the coronal extension of alveolar bone formation along the planed root; *Bone regeneration (area)*: area represented by new alveolar bone along the planed root; *Bone regeneration (density)*: ratio of mineralized bone matrix to total bone area; *Biomaterial density*: ratio of residual biomaterial to total bone area; *Root resorption*: combined linear heights of distinct resorption lacunae on the planed root; *Ankylosis*: combined linear heights of anky-
 lotic unions between new alveolar bone and the planed root

However, for calibration exercises, patient-based evaluations may focus on representative mean scores for the dentition (Lang et al. 2010).

6.6.1.1 Patient-Based Parameters

Running a probe along the gingival margin, a requirement for proper assessment of the oral hygiene level at the gingival margin (Plaque Index; Silness and Löe 1964) may remove small plaque deposits, and therefore, repeat assessment of the oral hygiene level is most likely to result in lower plaque scores (Lang et al. 2010; Shaw and Murray 1977; Birkeland and Jorkjend 1975). It also appears that plaque index scores cannot be reproduced within a 6-h period, and this is almost certainly due to the fact that the index requires that plaque should be removed with a probe (Shaw and Murray 1977). Similarly, repeated probing required for proper assessment of

gingival conditions utilizing any index system with bleeding as a criterion (GI; Löe and Silness 1963) may create gingival irritation and result in higher scores at subsequent examinations (Lang et al. 2010; Birkeland and Jorkjend 1975).

To reduce the number of passes within a dentition and still obtain a representative mean for a subject, half-mouth assessments or even evaluations of representative teeth for the dentition may be recommended for calibration exercises. For example, a mean score based on representative teeth of the mouth could be calculated for six patients with varying degrees of plaque and gingival health or disease. These six patients could be scored four times each, not consecutively, but in a random fashion. The first and the third scorings utilize exclusively the teeth proposed by Ramfjord (1959) as being representative of the dentition, while the second

and fourth examinations utilize the contralateral teeth of Ramfjord (1959) as being representative for the dentition. All evaluations involve three buccal aspects (mesio-buccal, buccal, and disto-buccal) and one lingual aspect (Lang et al. 2010).

Paired *t* tests may be used to assess systematic bias between the first and second subject-based mean scores or, in case of partial but representative recordings, between original and contralateral representative tooth mean scores. In case of multiple assessments, an analysis of variance to test the hypothesis of no difference between the assessments may be employed (Lang et al. 2010).

6.6.1.2 Site-Based Parameters

When considering the reproducibility of diagnosis, although a site-by-site analysis may appear initially to be more sensitive than using a total mouth score, there are certain disadvantages. First, if reproducibility is evaluated at sites using an index with only three possible scores, then the examiners could be expected to agree by chance one-third of the time. Secondly, when data for each site is analysed, scores of zero have been included in the calculation of reproducibility. This is in contrast to the FDI recommendation concerning the diagnostic reproducibility of caries, which excludes all sites diagnosed as sound. As reproducibility of diagnosis is partially dependent on the level of disease, the inclusion of many units free from disease could give a false assurance of diagnostic reproducibility. Thirdly, when the score for each site is used for computation, it is assumed that each site in the mouth contributed a value independently, whereas in reality, scores at the 24 designated sites are not truly independent but are related to the subject's oral hygiene pattern. Indeed, the main purpose of using carefully designated teeth rather than every tooth is to enable one to assess the gingival condition or plaque level of the whole mouth more quickly. However, using the raw scores of the whole mouth rather than

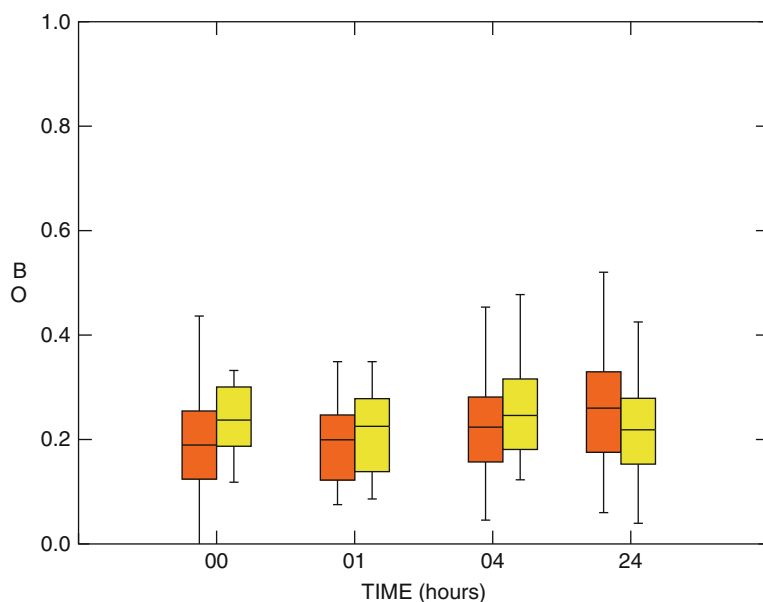
individual site scores does have one potential drawback: the site scores could be quite different on two separate occasions, but it is possible that the overall mouth scores might be identical (Shaw and Murray 1977).

To determine the reproducibility of single-site scores, the clinical protocols and the results of the calibration exercise for mean PII and GI may be used. However, the analysis will concentrate on the first and the third as well as on the second and fourth assessment of the sites. The reproducibility for the PII and the GI will be determined applying κ statistics or percentage agreement (Lang et al. 2010).

When assessing plaque, the scores of PII = 1 or 2 appear to be less reproducible than PII = 0 and 3 scores. This is owing that the latter scores are easily recognizable. At the second examination, the PII scores are usually slightly lower as it is likely that small amounts of plaque are removed during the first scoring as discussed above. Hence, a PII = 1 score may have been converted to a PII = 0 score (Lang et al. 2010).

A striking problem with gingivitis index is the lack of agreement as to the measurement criteria to be used and the evaluation standards to be employed. Examiner subjectivity as to what constitutes inflammation and the difficulty in accurately registering the related signs of gingival disease are among the major obstacles (Bessa Rebelo et al. 2011). When assessing gingival health, the lower GI scores (GI = 0 or 1) appear to be less reproducible than the higher GI scores of GI = 2 or 3. Bleeding of probing is the key feature of a GI = 2 score and is a more objective and easier to recognize feature than colour, texture, and swelling characterizing a GI = 1. At subsequent examinations, the GI site scores are usually higher as running the probe along the free gingival margin may disrupt the integrity of a slightly inflamed sulcus lining and, hence, convert a GI = 1 score to a GI = 2 score at a subsequent examination performed within 2 h (Fig. 6.4) (Lang et al. 2010; Birke-land and Jorkjend 1975; Müller and Barrieshi-Nusair 2005).

Fig. 6.4 Box plots showing proportions of sites with BOP at first (left) and second (right) probing with different time intervals (0, 1, 4, and 24 h) (Müller and Barrieshi-Nusair 2005. Reprinted with permission from Springer Science + Business Media)



Reports described in the literature have suggested that successive measures of gingival inflammation, evaluated by bleeding, performed by one or more examiners may not be reproducible (Feldman et al. 1982; Müller and Barrieshi-Nusair 2005). Consequently, the reproducibility of bleeding measurements has been a problem. This may have serious consequences for the conduction of clinical studies in which site-specific gingival bleeding is the main outcome variable (or an integral part of a more ‘precise’ index). As mentioned before, the probing procedure itself has a small detrimental effect as indicated by an increase in the proportion of bleeding sites after second probing among individuals and might last for more than 1 h. If repeatability of full-mouth bleeding scores is to be assessed 4 h after the first probing when this effect seems to be small and after 24 h, a small trial effect may be discernible. A certain degree of agreement of site-specific bleeding scores can only be observed when probing is repeated at once, whereas after 24 h, a different situation may be present. It can be concluded that reliability of gingival bleeding after probing in subjects with mild or moderate plaque-induced gingivitis can hardly be

determined in a replication study as adjusted associations between repeat BOP were weak in general, but strongest immediately after first probing (Müller and Barrieshi-Nusair 2005).

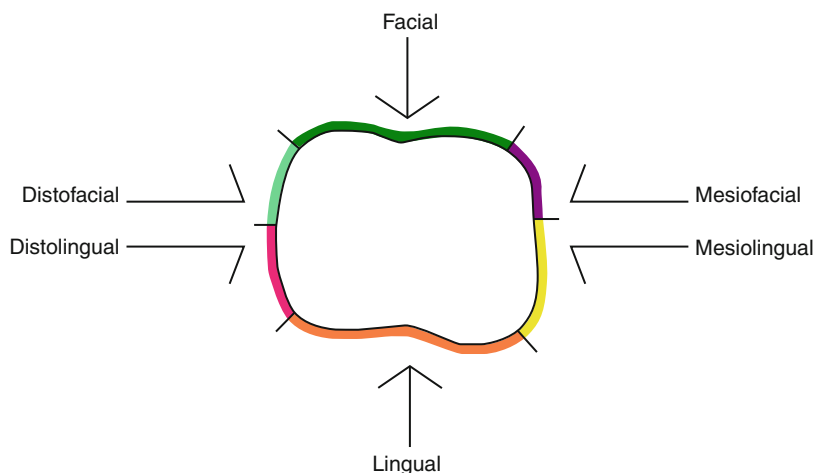
6.6.2 Continuous Parameters

The evaluation of periodontal pockets involves measuring probing depth and relating these measurements to a landmark, such as the cemento-enamel junction on the margin of a reconstruction in order to determine loss of connective tissue attachment (LA). Both parameters are expressed in millimetres and, hence, represent continuous parameters. Six aspects around a tooth or root are usually evaluated (Lang et al. 2010) (Fig. 6.5).

6.6.2.1 Patient-Based Parameters

For determining intra-examiner variability, it is proposed to examine either five patients five times, six patients four times, seven patients three times, or eight patients twice (Lang et al. 2010). Recently, Hill et al. (2006) calculated the number of calibration subjects based on a specified margin of error (half-width of the 95% confidence interval [CI]) of the percentage of agreement

Fig. 6.5 Periodontal probing on six sites per tooth



(exact and within 1 mm) for both intra- and inter-examiner reliability assessments. Specifically, they calculated the sample size (total number of tooth surfaces across all subjects), n , required to obtain a specified margin of error, ϖ , for a (population) proportion, π (in this case, percentage of agreement or agreement within 1 mm). The margin of error is given by $\varpi = 0.96 \times SE_{\text{cluster}}(p)$, where $SE_{\text{cluster}}(p) = [V_{\text{cluster}}(p)]^{1/2}$ is the SE of the estimated proportion under the clustered design; V_{cluster} is the variance of the statistic based on the clustered sampling design; and p , the sample proportion, is calculated here as the ratio of the total number of sites for which there is agreement to the total number of sites examined by both examiners and it follows $d_{\text{eff}} = V_{\text{cluster}} / V_{\text{independence}}$, where d_{eff} is the design effect; V_{cluster} is the variance of the statistic based on the clustered sampling design, and $V_{\text{independence}}$ is its variance based on a sample of the same number of independent observations. The number of calibration subjects $n = n_0 - d_{\text{eff}}$ (Eq. *), where n_0 is the sample size required to obtain the same level of precision in the estimation of π under an independent sampling design, and $n_0 = 1.96^2 \times \pi(1 - \pi) / \varpi^2$ (Eq. **).

Therefore, to find the sample size, n , needed to obtain a specified margin of error, ϖ , for a population proportion, π , under a cluster sampling design, Hill et al. (2006) recommended to:

1. Find the sample size, n_0 , assuming an independent sampling design using equation (Eq. **).
2. Multiply n_0 by the design effect to obtain the sample size, n , required under a cluster sampling design (Eq. *); the design effect is estimated for the corresponding sample proportion obtained from pilot data or from other published literature.
3. The resulting sample size, n , is the number of surfaces required to estimate percentage of agreement for each examiner pair under the cluster sampling design.

Again, it is reasonable to limit the number of passing to two when determining intra-examiner variability. Repeated probing of the same subject may severely compromise the well-being of that individual (Lang et al. 2010). Transient bacteraemia is common with manipulation of the teeth and periodontal tissues, and there is a wide variation in reported frequencies of bacteraemia in patients resulting from dental procedures: tooth extraction (10–100%), periodontal surgery (36–88%), scaling and root planing (8–80%), teeth cleaning (up to 40%), rubber dam matrix/wedge placement (9–32%), and endodontic procedures (up to 20%). Transient bacteraemia also occurs frequently during routine daily activities unrelated to a dental procedure, such as tooth brushing and flossing (20–68%), use of wooden toothpicks (20–40%), use of water irrigation devices (7–50%), and chewing food (7–51%) (Wilson et al. 2007). Significant results were also obtained regarding bacteraemia induced by

periodontal probing using a standard periodontal probe (Daly et al. 1997, 2001; Kinane et al. 2005) (Table 6.3). However, significant dental bacteraemia occurs whether or not there is discernible bleeding at the site of the operative procedure. Bleeding is a poor predictor of odontogenic bacteraemia (Roberts 1999).

The selection of a set of representative teeth and their contralateral counterparts as another set of representative teeth for the dentition has been advocated if patient-based outcomes are to be assessed. As mentioned for the non-parametric indices, the first and third scorings will consider exclusively the teeth proposed by Ramfjord (1959) as being representative of the dentition, while the second and fourth examinations, however, will consider only the contralateral teeth of Ramfjord (1959) as being representative of the dentition. The teeth are evaluated on three buccal and three lingual aspects (Lang et al. 2010).

If the mean values for PD and LA were generated with representative teeth and their contralateral counterparts, a Student's *t* test may be performed to verify that both groups of teeth would be representative of the whole dentition. Pending a positive outcome of no difference between the mean scores, the four scorings may be analysed by an analysis of variance with, expectedly, no difference between the mean scores for all assessments (Lang et al. 2010).

6.6.2.2 Site-Based Parameters

To determine the reproducibility of single-site measurements, the clinical protocol and the results of the calibration exercise for mean PD and LA may be used. However, the analysis will have to concentrate either on the first and the third or on the second and fourth assessments of sites (Lang et al. 2010).

The reproducibility of probe measurements has been shown to be dependent on a number of variables, which may be related to the pocket, the clinician, and the probe itself. These possible sources of error derived from the probe, angulation, and disease status have been shown to affect the accurate measurement of the true

pocket depth. Moreover, in case of repeated measurements by the same examiner or in cases where pocket depth measurements are performed by different examiners, accurate reproducibility of the measurement of the pocket depth may be influenced by these factors. This constitutes a major disadvantage of *in vivo* studies as it is not clear which of these factors are contributing to the variation. Furthermore, *in vivo* studies are seldom capable of assessing accuracy, as the true depth of the pocket is not known (Buduneli et al. 2004). A remedial scheme, the **option-3 scheme**, has been suggested in the literature. The scheme, called 'option-3' (Clark et al. 1993), uses two probings per site and evaluates how well the two measurements agree with each other. If the difference between the measurements exceeds a threshold value, a third measurement is taken. Two questions are eminent with respect to this procedure: What is the optimum threshold that triggers the use of option-3, and what is the appropriate estimate of the measurement for a site that needs option-3 adjustment? Namgung and Yang (1994) approached the first question and estimated the mean of the target site by averaging the closest two of the three values. The threshold was found to be proportional to the measurement error, ranging for attachment-level measurement is 1.2 mm $<|dl| < 2.8$ mm. A higher outlier rate must be anticipated in subjects with moderate to severe periodontitis (Hefti et al. 1997).

For ordinal variables (e.g. PD/LA within a 1-mm increment), percentage agreement remains an option, but treats non-agreements equivalently. Alternatives include reporting agreement within 1 mm or weighted κ (κ_w). The latter, κ_w , differentially weights non-agreements based on their relative magnitude while adjusting for chance agreement (Lang et al. 2010). The intra-examiner variability for PD and LA has been identified to lie between 0.36–0.41 and 0.54 mm, respectively. The calibration values for PD are normally slightly below than those for LA owing to the difficulties associated with identifying the cemento-enamel junction (Lang et al. 2010).

Table 6.3 Summary of studies on bacteremia associated with periodontal probing

Reference	Subjects	Sampling time	Identified species	Outcomes
Daly et al. (1997)	30 Patients with untreated periodontitis	Before and immediately after probing	Aerobic/facultative isolates (<i>Viridans streptococcus</i> spp. 45%, <i>Corynebacterium</i> spp. 15%, <i>coagulase-negative staphylococci</i> 10%, <i>Gemella</i> spp. 5%, <i>Streptococcus milleri</i> 5%), <i>Anaerobic isolates (Bacteroides</i> spp. 10%, <i>Desulfomonas</i> spp. 5%, <i>Peptostreptococcus</i> spp. 5%)	13 Patients (43%) exhibited bacteraemia of oral origin. No association was found between the severity of periodontitis, as determined by the deepest pockets recorded, and the occurrence of bacteraemia. The results indicate that periodontal probing can cause bacteraemia in patients with periodontitis
Daly et al. (2001)	20 Patients with adult periodontitis and 20 with chronic gingivitis	Prior to and immediately following periodontal probing	<i>Streptococcus</i> spp. were the most common isolates in both groups	Probing caused bacteremia of oral origin in eight (40%) of the periodontitis patients and two (10%) of the gingivitis patients. Compared with the gingivitis group, the odds ratio (OR) for bacteremia in the periodontitis group was 5.993 (95% CI 1.081–33.215). Bleeding on probing (OR 1.025, 95% CI 1.004–1.047) and mean probing depth per tooth (OR 1.444, 95% CI 1.055–1.977) were significantly associated with bacteremia. No significant correlations were found between bacteremia and age, number of teeth probed, smoking status, PI, and total probing depth
Kinane et al. (2005)	30 Volunteers with untreated periodontal disease	Sample 1: at baseline, Sample 2: following full-mouth periodontal probing depth (range 30 s to 1 min) Sample 3: at the second visit, 2 weeks later, Sample 4: after supervised 2-min toothbrushing (never more than 3 min) Sample 5: immediately following full-mouth ultrasonic scaling	Several organisms were detected: <i>Anaerobic Streptococci</i> , <i>Propionibacterium acnes</i> , <i>Neisseria pharyngis</i> , <i>Streptococcus viridans</i> , <i>Micrococcus</i> , <i>Staphylococcus albus</i> , <i>Hemophilus aphrophilus</i> , <i>Coag-ve intermedia</i> , <i>Actinomyces naeslundii</i> , <i>Gemella haemolysins</i> , <i>Streptococcus parasanguis</i> , <i>A. naeslundii</i> , <i>Eubacterium</i> sp., <i>Eubacterium limosum</i> , <i>A. Naeslundii</i> , <i>Streptococcus parasanguis</i> , <i>Propionibacterium acnes</i>	Using culture methods, the incidence of bacteraemias was as follows: following ultrasonic scaling (13%), periodontal probing (20%), and toothbrushing (3%). PCR analysis revealed bacteraemia incidences following ultrasonic scaling, periodontal probing, and toothbrushing of 23%, 16%, and 13%, respectively

References

- Aardal E, Holm AC. Cortisol in saliva-reference ranges and relation to cortisol in serum. *Eur J Clin Chem Clin Biochem.* 1995;33:927–32.
- Abbas F, Hart AA, Oosting J, van der Velden U. Effect of training and probing force on the reproducibility of pocket depth measurements. *J Periodontol Res.* 1982;17:226–34.
- Abrahamsson I, Soldini C. Probe penetration in periodontal and peri-implant tissues. An experimental study in the beagle dog. *Clin Oral Implants Res.* 2006;17:601–5.
- Albandar JM, Kingman A. Gingival recession, gingival bleeding, and dental calculus in adults 30 years of age and older in the United States, 1988–1994. *J Periodontol.* 1999;70:30–43.
- Albandar JM, Brunelle JA, Kingman A. Destructive periodontal disease in adults 30 years of age and older in the United States, 1988–1994. *J Periodontol.* 1999;70:13–29.
- Axtelius B, Edwardsson S, Theodorsson E, Svensäter G, Attström R. Presence of cortisol in gingival crevicular fluid. A pilot study. *J Clin Periodontol.* 1998;25:929–32.
- Barendregt DS, Van der Velden U, Reiker J, Loos BG. Clinical evaluation of tine shape of 3 periodontal probes using 2 probing forces. *J Clin Periodontol.* 1996;23:397–402.
- Bartko JJ. Measurement and reliability: statistical thinking considerations. *Schizophr Bull.* 1991;17:483–9.
- Bergmann A, Deinzer R. Daytime variations of interleukin-1 β in gingival crevicular fluid. *Eur J Oral Sci.* 2008;116:18–22.
- Bessa Rebelo MA, Adriana Corra de Queiroz A. Gingival indices: state of art. In: Panagakos FS, Davies RM, editors. *Gingival diseases – their aetiology, prevention and treatment.* InTech, São Paulo. 2011; p. 41–54.
- Birek P, McCulloch CA, Hardy V. Gingival attachment level measurements with an automated periodontal probe. *J Clin Periodontol.* 1987;14:472–7.
- Birkeland JM, Jorkjend L. The influence of examination of the assessment of the intra-examiner error by using the plaque and gingival index systems. *Community Dent Oral Epidemiol.* 1975;3:214–6.
- Blieden TM, Caton JG, Proskin HM, Stein SH, Wagener CJ. Examiner reliability for an invasive gingival bleeding index. *J Clin Periodontol.* 1992;19:262–7.
- Buduneli E, Aksoy O, Köse T, Atilla G. Accuracy and reproducibility of two manual periodontal probes. An in vitro study. *J Clin Periodontol.* 2004;31:815–9.
- Bulthuis HM, Barendregt DS, Timmerman MF, Loos BG, van der Velden U. Probe penetration in relation to the connective tissue attachment level: influence of tine shape and probing force. *J Clin Periodontol.* 1998;25:417–23.
- Caton J, Greenstein G, Polson AM. Depth of periodontal probe penetration related to clinical and histologic signs of gingival inflammation. *J Periodontol.* 1981;52:626–9.
- Chamberlain AD, Renvert S, Garrett S, Nilvéus R, Egelberg J. Significance of probing force for evaluation of healing following periodontal therapy. *J Clin Periodontol.* 1985;12:306–11.
- Cicchetti DV, Heavens JR. A computer program for determining the significance of the difference between pairs of independently derived values of kappa or weighted kappa. *Educ Psychol Meas.* 1981;41:189–93.
- Clark WB, Magnusson I, Namgung YY, Yang MCK. The strategy and advantage in use of an electronic probe for attachment measurement. *Adv Dent Res.* 1993;7:152–7.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37–46.
- Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70:213–20.
- Daly C, Mitchell D, Grossberg D, Highfield J, Stewart D. Bacteraemia caused by periodontal probing. *Aust Dent J.* 1997;42:77–80.
- Daly CG, Mitchell DH, Highfield JE, Grossberg DE, Stewart D. Bacteremia due to periodontal probing: a clinical and microbiological investigation. *J Periodontol.* 2001;72:210–4.
- Deas DE, Moritz AJ, Mealey BL, McDonnell HT, Powell CA. Clinical reliability of the “furcation arrow” as a diagnostic marker. *J Periodontol.* 2006;77:1436–41.
- Deinzer R, Mossanen BS, Herforth A. Methodological considerations in the assessment of gingival crevicular fluid volume. *J Clin Periodontol.* 2000;27:481–8.
- Dombret B, Matthijs S, Moradi Sabzevar M. Interexaminer reproducibility of ordinal and interval-scaled plaque indices. *J Clin Periodontol.* 2003;30:630–5.
- Eaton KA, Rimini FM, Zak E, Brookman DJ, Newman HN. The achievement and maintenance of inter-examiner consistency in the assessment of plaque and gingivitis during a multicentre study based in general dental practices. *J Clin Periodontol.* 1997;24:183–8.
- Egelberg J. Oral hygiene methods the scientific way (synopses of clinical studies). 1sted. Sweden: OdontoScience; 1999. p. 271–8.
- Eickholz P, Steinbrenner H, Lenhard M, Marquardt M, Holle R. Interexaminer reliability of the assessment of clinical furcation parameters as related to different probes. *Eur J Oral Sci.* 1999;107:2–8.
- Feldman RS, Douglass CW, Loftus ER, Kapur KK, Chauncey HH. Interexaminer agreement in the measurement of periodontal disease. *J Periodontol Res.* 1982;17:80–9.
- Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull.* 1971;76:378–82.
- Fleiss JL, Chilton NW. The measurement of interexaminer agreement on periodontal disease. *J Periodontol Res.* 1983;18:601–6.
- Fleiss JL, Gordon J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas.* 1973;33:613–9.
- Fleiss JL, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. *Psychol Bull.* 1969;72:323–7.
- Fleiss JL, Nee JCM, Landis JR. Large sample variance of kappa in the case of different sets of raters. *Psychol Bull.* 1979;86:974–7.

- Fleiss JL, Mann J, Paik M, Goultschin J, Chilton NW. A study of inter- and intra-examiner reliability of pocket depth and attachment level. *J Periodontol Res*. 1991;26:122–8.
- Gibbs CH, Hirschfeld JW, Lee JG, Low SB, Magnusson I, Thousand RR, Yerneni P, Clark WB. Description and clinical evaluation of a new computerized periodontal probe – the Florida probe. *J Clin Periodontol*. 1988;15:137–44.
- Goodson JM, Kondon N. Periodontal pocket depth measurements by fiber optic technology. *J Clin Dent*. 1988;1:35–8.
- Goulet JP, Clark GT. Clinical TMJ examination methods. *J Calif Dent Assoc*. 1990;18:25–33.
- Gröndahl K, Sundén S, Gröndahl HG. Inter- and intraobserver variability in radiographic bone level assessment at Brånemark fixtures. *Clin Oral Implants Res*. 1998;9:243–50.
- Grossi SG, Dunford RG, Ho A, Koch G, Machtei EE, Genco RJ. Sources of error for periodontal probing measurements. *J Periodontol Res*. 1996;31:330–6.
- Hansen AM, Garde AH, Persson R. Sources of biological and methodological variation in salivary cortisol and their impact on measurement among healthy adults: a review. *Scand J Clin Lab Invest*. 2008;68:448–58.
- Hausmann E. Radiographic and digital imaging in periodontal practice. *J Periodontol*. 2000;71:497–503.
- Hefti AF. Periodontal probing. *Crit Rev Oral Biol Med*. 1997;8(3):336–56.
- Hill EG, Slate EH, Wiegand RE, Grossi SG, Salinas CF. Study design for calibration of clinical examiners measuring periodontal parameters. *J Periodontol*. 2006;77:1129–41.
- Hou GL, Hung CC, Yang YS, Shieh TY, Tsai CC. Radiographic alveolar bone loss in untreated Taiwan Chinese subjects with adult periodontitis measured by the digital scanning radiographic image analysis method. *Dentomaxillofac Radiol*. 2003;32:104–8.
- http://en.wikipedia.org/wiki/Fleiss%27_kappa. Fleiss' kappa. Accessed on 22 Feb 2010.
- Jeffcoat MK. Current concepts in periodontal disease testing. *J Am Dent Assoc*. 1994;125:1071–8.
- Jeffcoat MK, Jeffcoat RL, Jens SC, Captain K. A new periodontal probe with automated cemento-enamel junction detection. *J Clin Periodontol*. 1986;13:276–80.
- Kawamura M, Fukuda S, Inoue C, Sasahara H, Iwamoto Y. The validity and reproducibility of an oral rating index as a measurement of gingival health care and oral hygiene level in adults. *J Clin Periodontol*. 2000;27:411–6.
- Khocht A, Chang KM. Clinical evaluation of electronic and manual constant force probes. *J Periodontol*. 1998;69:19–25.
- Kida IA, Aström AN, Strand GV, Masalu JR. Clinical and socio-behavioral correlates of tooth loss: a study of older adults in Tanzania. *BMC Oral Health*. 2006;6:5.
- Kinane DF, Riggio MP, Walker KF, MacKenzie D, Shearer B. Bacteraemia following periodontal procedures. *J Clin Periodontol*. 2005;32:708–13.
- Kingman A, Albandar JM. Methodological aspects of epidemiological studies of periodontal diseases. *Periodontol* 2000. 2002;29:11–30.
- Koo KT, Polimeni G, Albandar JM, Wikesjö UM. Periodontal repair in dogs: analysis of histometric assessments in the supraalveolar periodontal defect model. *J Periodontol*. 2004;75(12):1688–93.
- Kudielka BM, Hellhammer DH, Wüst S. Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology*. 2009;34:2–18.
- Lafzi A, Mohammadi AS, Eskandari A, Pourkhamneh S. Assessment of intra- and inter-examiner reproducibility of probing depth measurements with a manual periodontal probe. *JODDD*. 2007;1:19–25.
- Lang NP, Cullinan MP, Holborow DW, Heitz-Mayfield LJA. Examiner training: standardization and calibration in periodontal studies. In: Giannobile WV, Burt BA, Genco RJ, editors. *Clinical research in oral health*. 1st ed. Hoboken: Wiley-Blackwell; 2010. p. 159–75.
- Larsen C, Barendregt DS, Slot DE, Van der Velden U, Van der Weijden F. Probing pressure, a highly undervalued unit of measure in periodontal probing: a systematic review on its effect on probing pocket depth. *J Clin Periodontol*. 2009;36:315–22.
- Löe H, Silness J. Periodontal disease in pregnancy. 1. Prevalence and severity. *Acta Odontol Scand*. 1963;21:533–51.
- Magnusson I, Fuller WW, Heins PJ, Rau CF, Gibbs CH, Marks RG, Clark WB. Correlation between electronic and visual readings of pocket depths with a newly developed constant force probe. *J Clin Periodontol*. 1988;15:180–4.
- Mariath AA, Bressani AE, Haas AN, Araujo FB, Rösing CK. Professional flossing as a diagnostic method for gingivitis in the primary dentition. *Braz Oral Res*. 2008;22:316–21.
- Marks RG, Magnusson I, Taylor M, Clouser B, Maruniak J, Clark WB. Evaluation of reliability and reproducibility of dental indices. *J Clin Periodontol*. 1993;20:54–8.
- Mayfield L, Bratthall G, Attström R. Periodontal probe precision using 4 different periodontal probes. *J Clin Periodontol*. 1996;23:76–82.
- McClanahan SF, Bartizek RD, Biesbrock AR. Identification and consequences of distinct Löe-Silness gingival index examiner styles for the clinical assessment of gingivitis. *J Periodontol*. 2001;72:383–92.
- Mojon P, Chung J-P, Favre P, Budtz-Jorgensen E. Examiner agreement on periodontal indices during dental surveys of elders. *J Clin Periodontol*. 1996;23:56–9.
- Mombelli A, Mühle T, Brägger U, Lang NP, Bürgin WB. Comparison of periodontal and peri-implant probing by depth-force pattern analysis. *Clin Oral Implants Res*. 1997;8:448–54.
- Müller HP, Ulbrich M. Alveolar bone levels in adults as assessed on panoramic radiographs. (I) Prevalence, extent, and severity of even and angular bone loss. *Clin Oral Investig*. 2005;9(2):98–104.

- Müller HP, Barrieshi-Nusair KM. Gingival bleeding on repeat probing after different time intervals in plaque-induced gingivitis. *Clin Oral Investig*. 2005;9:278–83.
- Namung YY, Yang MC. Outlier reduction by an option-3 measurement scheme. *Biometrics*. 1994;50(1):173–82.
- Neto JB, Filho GR, Tramontina VA, Sallum EA, Nociti Jr FH, Sallum AW. Millimeter marks and probe tip diameter standardisation from commercially available periodontal probes. A comparative study. *J Int Acad Periodontol*. 2001;3:57–60.
- Nield-Gehrig JS. Fundamentals of periodontal instrumentation and advanced root instrumentation. Philadelphia: Lippincott Williams & Wilkins; 2007.
- Oakley C, Brunette DM. The use of diagnostic data in clinical dental practice. *Dent Clin North Am*. 2002;46:87–115.
- Osborn J, Stoltenberg J, Huso B, Aeppli D, Pihlstrom B. Comparison of measurement variability using a standard and constant force periodontal probe. *J Periodontol*. 1990;61:497–503.
- Osborn JB, Stoltenberg JL, Huso BA, Aeppli DM, Pihlstrom BL. Comparison of measurement variability in subjects with moderate periodontitis using a conventional and constant force periodontal probe. *J Periodontol*. 1992;63:283–9.
- Pecoraro M, Azadivatan-le N, Janal M, Khocht A. Comparison of observer reliability in assessing alveolar bone height on direct digital and conventional radiographs. *Dentomaxillofac Radiol*. 2005;34:279–84.
- Pierro VS, de Souza IP, Luiz RR, Barcelos R, Moraes RS. Reliability of two methods for measurement of alveolar bone level in children. *Dentomaxillofac Radiol*. 2008;37:34–9.
- Pihlstrom BL. Measurement of attachment level in clinical trials: probing methods. *J Periodontol*. 1992;63:1072–7.
- Polson AM. The research team, calibration, and quality assurance in clinical trials in periodontics. *Ann Periodontol*. 1997;2:75–82.
- Preshaw PM, Kupp L, Hefti AF, Mariotti A. Measurement of clinical attachment levels using a constant force periodontal probe modified to detect the cementoenamel junction. *J Clin Periodontol*. 1999;26:434–40.
- Quirynen M, Callens A, van Steenberghe D, Nys M. Clinical evaluation of a constant force electronic probe. *J Periodontol*. 1993;64:35–9.
- Rae G. The equivalence of multiple rater kappa statistics and intraclass correlation coefficients. *Educ Psychol Meas*. 1988;48:367–74.
- Ramfjord SP. Indices for prevalence and incidence of periodontal disease. *J Periodontol*. 1959;30:51–9.
- Reddy MS, Palcanis KG, Gears NC. A comparison of manual and controlled-force attachment-level measurements. *J Clin Periodontol*. 1997;24:920–6.
- Rise J, Tollefsen T. Reliability of plaque and periodontal measurements estimated by the internal consistency method. *Acta Odontol Scand*. 1984;42:293–6.
- Roberts GJ. Dentists are innocent! “Everyday” bacteremia is the real culprit: a review and assessment of the evidence that dental surgical procedures are a principal cause of bacterial endocarditis in children. *Pediatr Cardiol*. 1999;20:317–25.
- Sakka S, Al-ani Z, Kasioumis T, Worthington H, Coulthard P. Inter-examiner and intra-examiner reliability of the measurement of marginal bone loss around oral implants. *Implant Dent*. 2005;14:386–8.
- Seabra RC, Costa FO, Costa JE, Van Dyke T, Soares RV. Impact of clinical experience on the accuracy of probing depth measurements. *Quintessence Int*. 2008;39:559–65.
- Shaw L, Murray JJ. Diagnostic reproducibility of periodontal indices. *J Periodontal Res*. 1977;12:141–7.
- Shi XQ, Eklund I, Tronje G, Welander U, Stamatakis HC, Engström PE, Engström GN. Comparison of observer reliability in assessing alveolar bone changes from color-coded with subtraction radiographs. *Dentomaxillofac Radiol*. 1999;28:31–6.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–8.
- Silness J, Loe H. Periodontal disease in pregnancy. 2. Correlation between oral hygiene and periodontal condition. *Acta Odontol Scand*. 1964;22:121–35.
- Spolsky VW, Gornbein JA. Comparing measures of reliability for indices of gingivitis and plaque. *J Periodontol*. 1996;67:853–9.
- Suh YI, Lundgren T, Sigurdsson T, Riggs M, Crigger M. Probing bone level measurements for determination of the depths of class II furcation defects. *J Periodontol*. 2002;73:637–42.
- Teeuw WJ, Coelho L, Silva A, van der Palen CJNM, Lessmann FGJM, van der Velden U, Loos BG. Validation of a dental image analyzer tool to measure alveolar bone loss in periodontitis patients. *J Periodontal Res*. 2009;44:94–102.
- Valachovic RW, Douglass CW, Berkey CS, McNeil BJ, Chauncey HH. Examiner reliability in dental radiography. *J Dent Res*. 1986;65:432–6.
- Van der Zee E, Davies EH, Newman HN. Marking width, calibration from tip and tine diameter of periodontal probes. *J Clin Periodontol*. 1991;18:516–20.
- Verhoeven JW, Cune MS, de Putter C. Reliability of some clinical parameters of evaluation in implant dentistry. *J Oral Rehabil*. 2000;27:211–6.
- Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37:360–3.
- Walsh TF, Saxby MS. Inter- and intra-examiner variability using standard and constant force periodontal probes. *J Clin Periodontol*. 1989;16:140–3.
- Wilson W, Taubert KA, Gewitz M, Lockhart PB, Baddour LM, Levison M, Bolger A, Cabell CH, Takahashi M, Baltimore RS, Newburger JW, Strom BL, Tani LY, Gerber M, Bonow RO, Pallasch T, Shulman ST, Rowley AH, Burns JC, Ferrieri P, Gardner T, Goff D, Durack

- DT, American Heart Association Rheumatic Fever, Endocarditis and Kawasaki Disease Committee, Council on Cardiovascular Disease in the Young, Council on Clinical Cardiology, Council on Cardiovascular Surgery and Anesthesia, Quality of Care and Outcomes Research Interdisciplinary Working Group, American Dental Association. Prevention of infective endocarditis: guidelines from the American Heart Association: a guideline from the American Heart Association Rheumatic Fever, Endocarditis and Kawasaki Disease Committee, Council on Cardiovascular Disease in the Young, and the Council on Clinical Cardiology, Council on Cardiovascular Surgery and Anesthesia, and the Quality of Care and Outcomes Research Interdisciplinary Working Group. *J Am Dent Assoc.* 2007;138(739–45):747–60.
- Winter AA. Measurement of the millimeter markings of periodontal probes. *J Periodontol.* 1979;50:483–5.
- Wolf B, Von Bethlenfalvy E, Hassfeld S, Staehle HJ, Eickholz P. Reliability of assessing interproximal bone loss by digital radiography: intrabony defects. *J Clin Periodontol.* 2001;28:869–78.

The degree to which patients and research subjects act in accordance with the advice or instructions of their health-care provider or researchers is termed ‘adherence’ or ‘compliance’. These terms are commonly used interchangeably, albeit the former terminology may be preferred because of its more positive connotations. The challenges of non-adherence in clinical practice have considerable relevance for conducting research because of the corresponding impact of psychological, behavioural, and systems factors (Robiner 2005).

There are two primary types of adherence in research (Robiner 2005):

- **Regimen adherence** or **protocol adherence** refers to pursuing the assigned regimen consistently (e.g., keeping to schedule in taking medication at the specified dose). There may be multiple components of adherence (e.g., prescribed medication-taking schedules, maintenance of diet or exercise regimens, maintenance of oral health behaviours, self-monitoring, meeting objective biological targets, such as maintaining low glycosylated haemoglobin)
- **Follow-up adherence** refers to fulfilling the scheduled sequence of assessment measures within planned time windows (e.g., keeping follow-up clinic visits, undergoing laboratory tests) until reaching endpoint.

7.1 Compliance in Periodontal Antimicrobial Treatment

If antimicrobial agents are to be used in the treatment of periodontal infections, their efficacy will be dependent upon patient compliance. This is an important issue when the results of treatment are evaluated, because an unsatisfactory result could indicate either an ineffective agent or a non-compliant patient. An effective agent that is not used by the patient is of little value (Loesche 1999).

7.1.1 Systemic Agents

Regimen adherence problems can involve both errors of omission (e.g., forgetting to take a medication, delaying a dose, under-dosing), and commission (e.g., over-medicating, taking the wrong medication). Adherence problems can be random or non-random (e.g., covariates of illness severity, treatment effects, or length of time in study [i.e. study fatigue]), transitory or long term, or attributable to other factors with implications for data analysis and interpretation (Robiner 2005).

Patient noncompliance is a function of socioeconomic factors (Nagasawa et al. 1990), the clarity of the instructions (Stewart and Caranasos 1989), and the daily frequency at which the

medication should be taken (Cockburn et al. 1987; Eisen et al. 1990).

A regimen in which metronidazole is taken twice a day would be better than a regimen in which it is taken three times a day. Doxycycline, and especially azithromycin, has dosage schedules that would promote patient compliance (Loesche 1999).

7.1.2 Locally Delivered Agents

The development of local delivery vehicles solves the problem of patient compliance, since the vehicles are professionally placed, thereby negating any reliance upon the patient to take the medication. However, the duration of treatment effect, which could be considered as a manifestation of compliance, will vary according to whether these vehicles are retained for the duration of treatment. If these vehicles are lost, then the pocket was not exposed to the full dosage, and in a sense the patient could be considered as non-compliant (Loesche 1999).

This argument cannot be applied to the biodegradable vehicles, because if the vehicle is not present at 1 week, it could have been lost from the pocket as a result of those forces which eject fibres and films, or because it was lost via its intended biodegradation (Loesche 1999).

The chlorhexidine-containing chip reportedly is retained for 7–10 days, and pockets so treated could be considered compliant. The 25% metronidazole gel and the 2% minocycline gel may not remain *in situ* long enough for the treated pockets to be considered compliant (Loesche 1999).

7.2 Enhancing Adherence in Clinical Research

The adverse effects and costs of subject adherence for clinical research reveal a compelling need for more effective techniques for promoting, sustaining, and evaluating subject adherence through all research phases. The selection of subjects who understand study burdens and are willing to commit to regimens is a critical up-front

step in promoting research adherence (Shumaker et al. 2000; Robiner 2005). This requires a combination of effective screening to exclude subjects who might reasonably be identified as at risk for non-adherence based on histories of poor adherence with treatments, inconsistency with medical care (e.g., lateness or missed appointments), problematic communication, or ambivalence about participating. It also entails effective communication about trial participation (e.g., informed consent, accepting random allocation, importance of adherence) (Shumaker et al. 2000; Robiner 2005).

7.2.1 Pre-randomization Screening

If possible, investigators and staff should try to determine the participants' prior history with treatment interventions and their successes or failures with these behaviours. The best predictor of future adherence behaviour is past adherence behaviour, and potential participants are often candid about their former experiences with, for example, diets, drugs, and exercise programmes. Investigators should also attend to subtle and not so subtle cues during screening visits (e.g., rescheduling, lateness to visits, difficulty reaching participants by phone, participants' grimaces when staff describe study expectations, hesitations) (Shumaker et al. 2000).

Two pre-randomization approaches have been used primarily. The **'run-in'** is a behavioural methodology that places potential subjects on a prescribed regimen for a specified period to assess their fidelity in following a proxy regimen. The premise is that short-term adherence during pre-randomization can predict long-term adherence after enrollment. Run-ins are usually single-blind and may use placebos and pill counts. A second pre-randomization screening approach is **test-dosing**. Potential subjects are given a small dose of active agent for a few days prior to randomization. Test-dosing identifies those who may have difficulty tolerating adverse medication effects. It identifies candidates for exclusion based on their individual biological responses to specific agents (Robiner 2005).

Whereas pre-randomization screening approaches are appropriate for efficacy trials designed to evaluate whether agents are effective, they are not suitable for management trials that address the impact of interventions on broader populations. The benefit of pre-randomization screening is the potential to enhance the power of studies, thereby increasing the likelihood of finding bona fide group differences, while circumventing avoidable costs that ultimately would be associated with screened-out candidates' poor adherence (Robiner 2005).

7.2.2 Adherence Enhancement

The second primary approach for promoting adherence is to integrate adherence-enhancing strategies into research (see Table 7.1) (Robiner 2005).

7.2.3 Monitoring Adherence

A third critical component for clinical research is evaluating subjects' adherence to assigned regimens. Several methodologies have been employed. In pharmacological studies, markers (e.g., inert molecules, radioactive substances, stable isotopes, pharmacological substances with relatively long half-lives, such as low dose phenobarbital; digoxin) may be used, coformulated with study preparations (including placebo) to quantify exposure to agent. Another approach incorporates monitoring systems such as the Medication Event Monitor System (MEMS) marketed by the Aprax division of Advanced Analytical Research on Drug Exposure (AARDEX), using microchips embedded in drug vial caps to identify when vials are opened. The approach is predicated on the presumption that

Table 7.1 Adherence enhancement strategies in clinical research (Robiner 2005) (Reprinted with permission from Elsevier)

Recruitment
Spend adequate time providing informed consent and getting to know subject
Explain the disorder, protocol, medication/device, side effects, interactions
Explain participants' roles, need for and importance of randomization and adherence
Explore, motivation, expectations, views of research, tolerance of randomization, history of research participation; history of discipline in self-care
Set clear, mutually acceptable goals and assess understanding
Use adherence run-in, or other techniques prior to randomization to screen for adherence and exclude potential poor adherers
Balance staff objectives in recruitment between enrolling largest number of subjects possible and limiting enrollment to those with reasonable likelihood of adhering to allow study to provide interpretable results
Health care
Promote contact with other treating professionals and garner their support for subject to participate
Provide health-related and research-related information
Promote referrals as necessary to other health professionals to support optimal health care, including mental health support
Address emotional aspects of disease and research participation
Social aspects
Promote positive, collaborative relationships between subjects and members of research team
Provide social support
Maintain frequent phone contact, email, clinic visits
Send birthday and anniversary cards, trial newsletter
Provide positive feedback for regimen and follow-up adherence
Social reinforcement and other reinforcers
Involve family members, or other support network, to promote involvement in research when necessary; consider family therapy referral

(continued)

Table 7.1 (continued)**Regimen characteristics**

Goal setting: develop clear and realistic expectations

Streamline protocol

- Minimize dosing frequency, number of pills, risk of side effects

- Minimize uncomfortable and unpleasant aspects of protocol

- Ensure materials are readable, conveniently packaged

Tailoring: develop regimen that can be realistically integrated with patients' other daily activities

Optimize clinical regimen to minimize adverse effects

Reduce negative consequences associated with regimen

Enhance health-care access and promote treatment of comorbid conditions and addressing psychosocial factors

Logistical support

Schedule appointments at convenient times and in convenient locations; minimize waiting times

Provide free parking, babysitting, support with transportation and lodging (if needed)

Provide compensation for missed work, travel reimbursement

Promote access to relevant technology (e.g., use of Internet, email)

Provide reminders of regimen and importance of sustaining adherence throughout

Adherence

Dedicate resources to promoting adherence (e.g., incorporate staff and consultants with expertise in adherence, encourage staff professional development, such as negotiating around adherence, understanding motivations for and types of non-adherence, as well as ethical issues in research and adherence)

Devise adherence plan as part of study design, including protocol for addressing non-adherence

Promote collaboration between subjects and research staff

Provide feedback about how well subjects are adhering to protocol or achieving target goals whenever possible

Promote candid, non-judgmental discussion of adherence, including barriers, facilitators, and personal challenges

Anticipate adherence challenges and address them proactively (teach skills necessary for to organize behaviours that underlie adherence, time management, and stress management)

Monitor adherence (e.g., pill counts, Medication Event Monitoring System (MEMS) caps, journal, logs, assays, use of tracer substances, self-monitoring and self-report, collateral reports); direct observation (i.e. supervised therapy/dispensing)

Use multiple methods of assessing adherence, including, but not limited, to self-report

Incorporate behavioural techniques and problem-solving to enhance adherence

- Practical strategies to promote consistent pill-taking (e.g., pairing medication-taking with routine life events)

- Use prompts (e.g., pill boxes, reminder letters, telephone calls, emails, Internet)

- Provide positive reinforcement (i.e. social reinforcement, incentives) for good adherence whenever possible (e.g., at clinic visits, inter-visit communications)

- Model adherent behaviour (e.g., show subject other subjects successfully following regimen)

- Intervene early and as often as necessary when adherence problems emerge (e.g., call if an appointment is missed); discuss barriers to adherence

- Increase frequency of visits if adherence becomes problematic

- Decrease visit frequency if adherence is sustained

- Use behavioural contracts when necessary

Address factors, as possible, that indirectly may affect adherence (e.g., refer for mental health treatment or social services if depression appears to be a factor undermining adherence)

Systems

Promote research on enhancing adherence

Increase funding within clinical research awards to support adherence-enhancing efforts

medication is taken only through the monitored vial and that each vial opening indicates medication was actually taken (Robiner 2005).

Participants who have become poor adherers, non-adherers, and/or who have dropped out of the study completely represents a major challenge to randomized controlled trials and underscores the importance of non-adherence prevention strategies described earlier. Negotiating and contracting skills are crucial. People's lives change, and factors that may make adherence perceived to be impossible at a particular life stage may change with time. Thus, it is always best to maintain at least a minimum level of contact with a participant so that the door is left open for recontact, and, possibly, restarting the study protocol. This means that staff must also be sensitive to how hard to push at any given time and when to let go, at least temporarily, so that future opportunities to reengage a participant are available (Shumaker et al. 2000).

Finally, participants have a right to drop out of a study or to stop adhering to a study protocol. This fact must always be salient to investigators and staff. Though we may work with participants to maintain their active and full participation in a

randomized controlled trial, we must also know where to draw the line; participant autonomy and full rights are essential (Shumaker et al. 2000).

References

- Cockburn J, Gibberd RW, Reid AL, Sanson-Fisher RW. Determinants of non-compliance with short term antibiotic regimens. *Br Med J (Clin Res Ed)*. 1987;295:814–8.
- Eisen SA, Miller DK, Woodward RS, Spitznagel E, Przybeck TR. The effect of prescribed daily dose frequency on patient medication compliance. *Arch Intern Med*. 1990;150:1881–4.
- Loesche WJ. The antimicrobial treatment of periodontal disease: changing the treatment paradigm. *Crit Rev Oral Biol Med*. 1999;10:245–75.
- Nagasawa M, Smith MC, Barnes Jr JH, Fincham JE. Meta-analysis of correlates of diabetes patients' compliance with prescribed medications. *Diabetes Educ*. 1990;16:192–200.
- Robiner WN. Enhancing adherence in clinical research. *Contemp Clin Trials*. 2005;26:59–77.
- Shumaker SA, Dugan E, Bowen DJ. Enhancing adherence in randomized controlled clinical trials. *Control Clin Trials*. 2000;21:226S–32.
- Stewart RB, Caranasos GJ. Medication compliance in the elderly. *Med Clin North Am*. 1989;73:1551–63.

8.1 Introduction

Health care professionals are increasingly required to base their practice on the best available evidence. After a literature search on a specific clinical question, many articles may be retrieved. The quality of the studies may be variable, and the individual studies might have produced conflicting results. It is therefore important that health care decisions are not based solely on one or two studies without account being taken of the whole range of research information available on that topic (Akobeng 2005).

Systematic reviews have rapidly gained an important place in aiding clinical decision-making in medicine, although dentistry has been a little slower to adopt this approach. Systematic reviews are themselves considered a research activity, although the data are derived from primary studies in the area of interest rather than from direct experimentation (Needleman 2002).

8.1.1 What Is a Narrative Review and How Is It Different from a Systematic Review?

Traditional, narrative reviews, usually written by experts, are qualitative summaries of evidence on a given topic. Typically, they involve informal, subjective methods to collect and interpret studies, and tend to selectively cite literature that reinforces preconceived notions. Narrative reviews often do not explicitly describe how the reviewers

searched, selected, and appraised the quality of studies. In contrast, a systematic review includes a comprehensive, exhaustive search for primary studies on a focused clinical question, selection of studies using clear and reproducible eligibility criteria, critical appraisal of studies for quality, and synthesis of results according to a predetermined and explicit method (Pai et al. 2004).

To further understand systematic reviews, the differences compared with a traditional review are detailed in **Table 8.1** (Needleman 2002).

8.1.2 What Is a Systematic Review?

In contrast to a narrative review, a systematic review is a form of research that provides a summary of medical reports on a specific clinical question, using explicit methods to search, critically appraise, and synthesize the world literature systematically. It is particularly useful in bringing together a number of separately conducted studies, sometimes with conflicting findings, and synthesizing their results (Akobeng 2005).

To help to understand systematic reviews, it is helpful to make an analogy with clinical trials (**Table 8.2**). Part of the reason for the similarity in design is the importance attached in both types of research to transparency and minimizing bias (Needleman 2002).

The description of systematic reviews as providing the highest level of evidence is widespread but also raises expectations that may or may not be fulfilled. A realistic understanding of what a

Table 8.1 Comparison between a systematic and a traditional review (Needleman 2002) (Reprinted with permission from John Wiley & Sons, Inc.)

Systematic review	Traditional review
Method of review stated	Method not stated
Answers focused question and narrow in scope	No real question and often very broad in scope
All relevant studies located	Biased selection of studies
Quality of research systematically appraised	Results of studies often quoted without appraisal of quality
May combine results statistically in a meta-analysis. Conclusions based on the data derived	No real explanation for reviewer's conclusions
Results may illustrate the true level of heterogeneity of data allowing quantification of uncertainty	Results often presented in black and white terms without indicating the uncertainty or variability in results

Table 8.2 Analogy between a systematic review and the design of a clinical trial (Needleman 2002) (Reprinted with permission from John Wiley & Sons, Inc.)

Clinical trial	Systematic review
Based on stated hypothesis	Based on stated focused question
Pre-stated protocol specifying: Patient recruitment Patient inclusion/exclusion criteria Interventions Outcome measures to be assessed Data analysis	Pre-stated protocol specifying: Search strategy Study inclusion/exclusion criteria Intervention/exposure of interest Outcome measures to be assessed Data analysis

systematic review can provide is important for the appropriate use of this type of evidence (Table 8.3) (Needleman et al. 2005).

8.1.3 What Is a Meta-Analysis?

A meta-analysis is the statistical pooling of data across studies to generate summary (pooled) estimates of effects. The term ‘effect’ refers to any measure of association between exposure and outcome (e.g. odds ratio). A meta-analysis is usually the final step in a systematic review. All meta-analyses should ideally start with an unbiased systematic review that incorporates articles chosen using predetermined inclusion criteria. If the data extracted from these studies meet certain requirements (the most important being a high level of homogeneity of effect measures across studies), then the data can be combined using meta-analysis. However, if the effect measures

are found to be heterogeneous, then it is still acceptable to present the work as a systematic review and not perform meta-analysis, or use statistical methods that can account for the **heterogeneity**. Indeed, there are situations when a meta-analysis is clearly inappropriate. Therefore, meta-analyses and systematic reviews are not synonymous. Ideally, a meta-analysis should be performed as part of a systematic review. In practice, meta-analyses are sometimes done without an initial systematic review. Within the set of meta-analyses, the investigators will sometimes choose to go beyond the analyses of published studies, contact authors of the primary studies for data on individual patients in their studies, and then combine the raw data. This is called an **individual patient data (IPD) meta-analysis** (Pai et al. 2004).

Unfortunately, the vocabulary attendant to the process of meta-analysis is intimidating. (The various terms are defined in Table 8.4).

Table 8.3 The potential of systematic reviews (Needleman et al. 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

What a high quality systematic review can do

- Find and summarize all available studies.
 - A comprehensive search will identify all relevant studies up to the point of the date of the completion of the search.

This should give the reader greater confidence that bias in selecting studies has been minimized.

- Provide an objective assessment of the quality or research and in particular the degree of protection from bias within the original studies.
 - Components of methodological quality that have evidence of affecting bias can be evaluated. It may be much harder to evaluate the impact of other “quality” issues if there is no consensus on how to measure them. An example could be the quality of the treatment provided.
- Estimate research effects across multiple studies with meta-analysis.
 - Meta-analysis is valid only if studies are similar in their research question and design.
 - Meta-analysis can estimate uncertainty and precision of the effect.
 - Meta-analysis may generate hypotheses for differential effects across subgroups of the population tested.
- If the effect is consistent across multiple studies (with small differences in design), then it may more readily possible to generalize the results to clinical practice than the results from a single study.
- Overcome limitations of underpowered studies in detecting a true difference if such a true difference really exists.

What a high quality systematic review cannot do

- It cannot be used in isolation to dictate clinical practice.
- It is a synthesis of available research and must be used in context with clinical judgement and patient preference.
- Produce strong conclusions if the research base is weak in quality.
 - The value of the review will then be to present a comprehensive objective summary of the strength of the data and to identify the design of research to answer important gaps.
- Overcome limitations of narrowly designed clinical research.
 - If the clinical studies only investigate the effect of an intervention in highly selected individuals, the conclusions cannot be generalized outside of these conditions. However, the objective communication of any limitations in the research base will help to set the degree of uncertainty and indicate the priorities for future research.
- Exclude relevant studies.
 - Although the majority of hits from the search will be excluded, this is due to the deliberate strategy of achieving high sensitivity (likelihood of finding all relevant studies) but low precision (likelihood of only finding relevant studies).

Therefore, it is common to find that more than 90% of the search records are totally irrelevant to the question and must be excluded. The alternative approach of aiming for high precision also carries a high risk of missing relevant studies, although it will appear as if few studies are being excluded.

- Be a miracle research design.
 - All research has strengths and limitations/weaknesses. Systematic reviews are no different from other research designs in this respect.

Table 8.4 Definition of terms (Koretz 2002) (Reprinted with permission from Wolters Kluwer Health)

Term	Meaning
Absolute risk difference	The arithmetical difference between the rate of the event in question in the treated group and the rate of that event in the control group.
Confidence interval	If the experiment were to be repeated a 100 times, the ‘n’% (typically 95%) confidence interval quantitates the limits between which the observed result would fall ‘n’ of those times. (In other words, the n% confidence interval means that there is an n% likelihood that the true answer lies somewhere within that range.) In general, a meaningful difference is present if the 95% confidence interval does not overlap the line of no effect (0.0 for absolute risk differences or weighted/standardized mean differences, 1.0 for relative risks or odds ratios).
Continuous variable	A result that is not discrete (dichotomous).

(continued)

Table 8.4 (continued)

Term	Meaning
Data combination	The act of adding together observed results from two or more studies.
Data pooling	Data combination employing simple addition of all of the results without any weighting.
Dichotomous variable	A result that is either present or absent. (Categorical variables are results that can be classified as belonging to one of a limited number of possibilities; if that limited number is two, categorical and dichotomous have identical meanings.)
Fixed effects model	The mathematical model for meta-analysis that is used if the trials are believed to be homogeneous. (The model only employs intra-trial variances in the weighting system.)
Forest plot	The pictorial presentation of a meta-analysis.
Funnel plot analysis	A technique to detect publication bias. (The size of the effect is graphed against the size of the study. If no publication bias is present, the small trials should be equally distributed on both sides of the largest one or two studies [forming a bell-shaped curve or funnel].)
Heterogeneity	Studies that are unlike enough to cause concern about the appropriateness of employing data combination (or at least to result in modifications of the analytic procedures that are employed).
Homogeneity	Studies that are alike enough to permit data combination.
Line of no effect	Vertical line in Forest plot that represents equivalency in outcome for the two treatment arms (1.0 for relative risk ratio or odds ratio; 0.0 for absolute risk difference or weighted/standardized mean difference).
Meta-analysis	Data combination employing a weighting process for the individual studies such that those with less variance are given more emphasis.
Narrative review	A putative expert's overview of a topic in which no formal rules or standards need be used in its preparation.
Number needed to treat	The number of patients that would have to be given the intervention under study in order to achieve or avoid one event (calculated by dividing 100 by the absolute risk difference expressed as a percentage). (If there is a 5% difference favouring the treated group, one would have to treat 20 patients to obtain one favourable outcome.)
Odds ratio	The odds of a particular event occurring in the intervention group divided by the odds of it occurring in the control group. (The odds ratio is not easy to grasp intuitively; when the rates of the event are small in each group, the odds ratio approximates the relative risk ratio.)
Power	The likelihood that a particular difference will be statistically demonstrated given the number of individuals enrolled in a trial and the anticipated event rates in the interventional and control groups.
Precision	An expression of how close the observed result is to the true situation. (A high precision is reflected by a narrow confidence interval.)
Publication bias	The tendency for dramatic or significant results to be submitted to, and accepted by, medical journals.
Q-test	A statistical estimate of the presence of homogeneity.
Quality of the study	An evaluation of the scientific rigour (elements that reduce the influence of bias) of the study. (High quality indicates the presence of good rigour.)
Quality score	A numerical estimate of the quality of the study.
Quasirandomized	A method of allocation whereby the investigator or subject will know into which study arm assignment will occur prior to the subject being enrolled (e.g. assignment made by patient identification number or day of the week).
Random effects model	The mathematical model for meta-analysis that is used if the trials are believed to be heterogeneous. (The model employs both intra-trial and inter-trial variances in the weighting system.)
Relative risk ratio	The incidence of the event in the treated group divided by the incidence of the event in the control group. (For example, a risk ratio of 0.75 means that the event occurred only three-fourths as often in the treated group as in the control group.)

Table 8.4 (continued)

Term	Meaning
Sensitivity analysis	Confining the meta-analysis to a subgroup of trials that share a certain feature. (An a priori decision is made regarding what the important components of the heterogeneity are likely to be and separate analyses are performed that only include the trials that share these components.)
Standardized mean difference	The meta-analysis summary estimate for studies in which the outcome variable is continuous but not comparable between the studies (e.g. expressing the results of a test as a percentage of the upper limit of normal).
Subgroup analysis	See sensitivity analysis.
Summary estimate	The result of the meta-analytic calculation.
Systematic review	A structured effort to obtain the answer to a specific question by employing pre-planned literature searches, data abstraction, and analyses.
Type I error	A study describing a ‘significant difference’ in which the observed difference is truly due to chance.
Type II error	A study failing to see a difference that truly exists.
Variable	See continuous variable, dichotomous variable.
Variance	A statistical estimate indicating how far a particular observation is from the truth. (It reflects how close the individual observations are to each other, using the standard deviation or confidence interval.) (Although other factors play into the calculation of variance, larger trials usually have less variance.)
Weighted mean difference	The meta-analysis summary estimate for studies in which the outcome variable is continuous.
Weighting	The degree of emphasis that will be given to each study in the meta-analysis. (It is inversely proportional to the variance.)

8.2 Steps in Conducting a Systematic Review

The steps in a systematic review are illustrated in **Fig. 8.1**.

8.2.1 Structuring the Search for Information and Evidence: The PICO Method

8.2.1.1 Asking a Clinical Question

Clinical question formulation has been posited to be an important skill for physicians. Sackett’s widely read textbook on evidence-based medicine (Sackett 2000) provides a list of seven reasons for this, although this book acknowledges that there are no controlled trials that demonstrate the relationship between question formulation and better clinical care. Two of the reasons given seem especially relevant: Well-formulated questions help us, as physicians, to focus our learning time on learning needs that are directly relevant to our patients’

clinical needs, and well-formulated questions help us communicate more clearly with our colleagues (Bergus and Emerson 2005).

A clinical question is generally defined as a question pertaining to a health care provider’s management of a patient or population of patients that is answerable through print or electronic resources. **Clinical questions** can be divided into two general categories: background and foreground questions.

Background clinical questions can usually be answered by information from textbooks, websites, and hospital or office information resources such as patient history files. Background questions typically focus on general information needed or a specific concept (an intervention, an aspect of a disease or disorder, the determination of possible therapies). Background questions generally begin with a question root such as who, what, when, why, or how. A good background question might include what are effective massage or manipulative therapies for carpal tunnel syndrome (Sullivan 2008).

Foreground clinical questions usually seek to find relevant, sometimes individualized, evidence from primary research publications or sec-

ondary, synthesized publications built from primary research papers. Foreground questions typically include three to four component terms or keywords focused on the patient, intervention, comparison to the intervention, and outcome desired (Sullivan 2008).

8.2.1.2 PICO Question

Successful construction of Evidence-based periodontology questions represents an acquired rather

than an innate skill that needs mentoring and review for success. For support in structuring a foreground question, an acronym exists to act as a mnemonic for the necessary components of the question. The acronym, PICO, stands for patient, intervention, comparison, and outcome. Using **PICO as an anatomical question** framework helps assure the clinician that the question established is a foreground question. Another benefit to structuring a question in the PICO format is that the components of the question can assist in formulating search strategies for evidence, thus promoting evidence-based decisions (Larue et al. 2009; Sullivan 2008; Schardt et al. 2007; Bragge 2010):

Architecture of a focused question

Patient: Defining the patient characteristics is essential. A specific, narrow definition will provide very applicable evidence for that particular patient but may limit the evidence too much so that important evidence is excluded from search results. Consider keywords and phrases that will allow a health care provider imagine the patient in front of you (Sullivan 2008).

Intervention: An intervention component may be broad or narrow. When seeking ‘best evidence’, several interventions may be specified in separate foreground questions. Broad phrases (‘what is most effective?’) often lead to background questions. A specific intervention should suggest something that will influence the desired outcome (Sullivan 2008).

Comparison: The comparison component often is the ‘second half’ of the intervention component. In therapy questions, intervention A might be compared to a well-known or standard therapy. For diagnosis questions, the comparison is often the ‘gold standard’ diagnostic. Prognosis

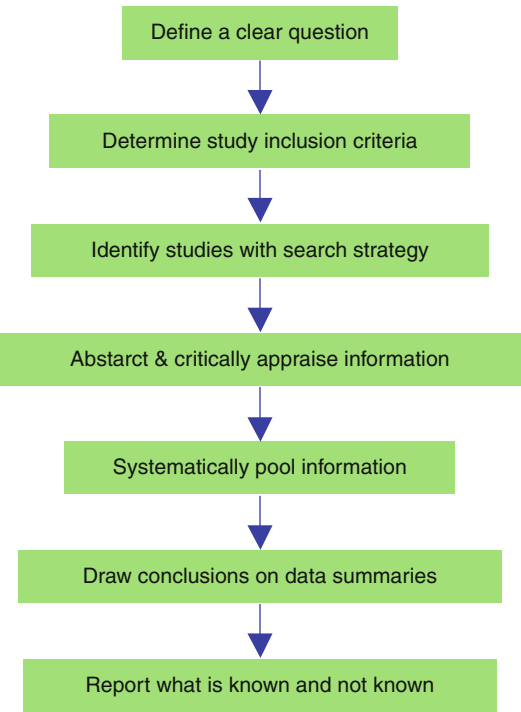


Fig. 8.1 Stages in systematic review process (Needleman 2002. Reprinted with permission from John Wiley & Sons, Inc.)

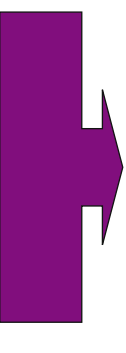
	P	Patient problem:	Who or What?
	I	Intervention:	How?
	C	Comparison:	What is the main alternative? (If appropriate)
	O	Outcome:	What are you trying to accomplish, measure, improve, effect?

Table 8.5 Type of evidence to seek when answering a clinical question (Koretz 2002) (Reprinted with permission from Wolters Kluwer Health)

Question relates to	Best evidence
Prognosis of disease	Natural history studies
Etiology of disease	Epidemiologic studies ^a
Utility of diagnostic tests	Sensitivity/specificity/predictive values
Intervention in disease	Randomized trials

^aRandomized trials that demonstrate that the treatment of the proposed etiology cures the disease are the strongest type of evidence, but such information is usually not available

or aetiology questions may include a factor which may affect the patient population in some way (Sullivan 2008).

Outcome: The last component of the PICO format is the outcome. The ‘outcome’ is what the clinician hopes to accomplish. An outcome should be patient oriented (taking patient values, expectations, preferences, and priorities into consideration), definable, measurable, and clinically relevant. In many cases, there will be more than one relevant, important outcome that depends on what aspect of care is needed or desired (Sullivan 2008).

PICO was first developed by Richardson et al. (1995) and later applied to numerous disciplines including physical therapy (Scherer and Smith 2002), occupational therapy (Law 2002), and various subspecialties within medicine (e.g. Armstrong 1999).

8.2.1.3 PICOTT Question

The PICO framework can be expanded to **PICOTT**, adding information about the type of question being asked (therapy, diagnosis, prognosis, harm, etc.) and the best type of study design for that particular question. Using this framework helps the clinician articulate the important parts of the clinical question most applicable to the patient and facilitates the searching process by identifying the key concepts for an effective search strategy (Scharadt et al. 2007; Wilczynski et al. 2002; Bergus and Emerson 2005).

Once a specific question is identified, a determination is made regarding what kind of research will best answer it (Table 8.5). An explicit literature search strategy is planned (usually involving at least a computer search) including criteria for using or not using each identified paper. A data

abstraction sheet is prepared with spaces for each piece of information that will be sought (including elements assessing the quality of the study). An a priori decision is made regarding data analysis, including a decision about data combination (the role of meta-analysis). Sensitivity (subgroup) analysis is planned to evaluate heterogeneity. The process is then undertaken. If pieces of information are missing, the original investigators can be contacted. At the end, the conclusions are the consequences of the analyses. Whatever data were found to be inadequate become the future research needs (Koretz 2002).

8.2.1.4 So How Does It Work in Practice?

Miller and Miller (2010) illustrate the use of the evidence-based process in clinical practice to answer questions that arose when a patient presented with dental trauma. The patient had avulsion of the maxillary right central incisor and lateral luxation and alveolar bone fracture partially encasing the roots of the maxillary left central and lateral incisors. At the emergency treatment, the dentist replanted the teeth but had two questions regarding the impending treatment for the patient regarding the optimal timing of root canal therapy and splint duration that would result in the best outcome and prognosis for healing. When searching for evidence, the PICO question guides the search. The articles that were selected as relevant research included each aspect of the PICO question. Inclusion criteria included the following: The patient population studied had to have replanted avulsed or luxated teeth; the research studied the intervention for each of the two PICO questions, pulp extirpation and splint duration, respectively; and measured at least one of the outcomes of tooth integration, functional periodontal healing, or the levels of resorption or ankylosis (Table 8.6). The highest level of evidence for this prognosis question is a meta-analysis or systematic review of cohort studies. Meta-analysis is a systematic review that uses the statistical process that combines the data from multiple individual studies into a single pooled estimate or analysis. A systematic review is a critical assessment and evaluation of two or more primary research studies that have investigated the same specific phenomenon or question and

Table 8.6 Search terms for each PICO question (Miller and Miller 2010) (Reprinted with permission from Elsevier)

PICO question 1 search terms	PICO component	PICO question 2 search terms
Tooth avulsion (MeSH) OR Tooth replantation (MeSH)	P	Tooth avulsion (MeSH) OR Tooth replantation (MeSH)
Pulp extirpation OR Root canal therapy (MeSH)	I	I Splints (MeSH)
(Same intervention as above, however timing is the real comparison so that is the factor in the final article selection)	C	(Same intervention as above, however timing is the real comparison so that is the factor in the final article selection)
Tooth integration OR Functional periodontal healing OR root resorption (MeSH) OR tooth ankylosis (MeSH)	O These terms were used as inclusion criteria and were not used when searching PubMed because only a few number of systematic reviews and guidelines were found just using the P, I, C terms	Tooth integration OR Functional periodontal healing OR root resorption (MeSH) OR tooth ankylosis (MeSH)

uses explicit predefined criteria for article retrieval, assessment, and synthesis of evidence from the individual studies to reduce bias. In the event that a systematic review or meta-analysis of cohort studies was not found, individual cohort studies are the next highest level of evidence. A cohort study is defined as a prospective investigation of the factors that may cause a disorder in which a group of individuals with a common characteristic or set of characteristics are exposed to a cause compared with a group that is not exposed (Miller and Miller 2010).

Use of Boolean operators represented by the connector terms AND, OR, and NOT. These terms allow for combinations of descriptors that will be used in the search, with AND for a restrictive combination, OR for an additive combination, and NOT for an excluding combination. Combination of components of the PICO strategy for the finalization of the search strategy: after the selection of the search terms and use of Booleans operators for each of the four components of the PICO strategy, these must be inter-related in the following final strategy: (P) AND (I) AND (C) AND (O). Such final strategy must be inserted in the search box existent in the databases so that evidence is located by means of a bibliographic search (Santos et al. 2007).

8.2.1.5 Evaluating the Impact of Formulated Questions

In a recent communication to the evidence-based-health discussion list (August 2005), Scott

Richardson, an early proponent of question formulation, outlines seven potential benefits from the question formulation process (Booth 2006). Although research into question formulation within our own professional context is much more immature, it is nevertheless possible to extrapolate that these benefits, albeit slightly reworded, also exist for our own evidence-based practice (Booth 2006):

1. To use our scarce learning time on evidence that directly relates to our users’ needs
2. To optimize our scarce learning time on evidence that directly relates to our own learning needs
3. To suggest high-yield search strategies
4. To suggest the forms that useful answers might take
5. To improve articulation and communication of problems with colleagues
6. When teaching, to help our learners understand the content of what we teach, while modelling a useful skill for lifelong learning
7. To build both our knowledge base and our learning skills and to reward (rather than punish) curiosity

Formulating the question is fundamental to evidence-based practice, irrespective of the discipline involved. Question formulation, and indeed question answering, is a key competency for our profession. Our practice may be informed both by research within information science and by wider developments in evidence-based practice (Booth 2006).

8.2.2 Search and Inclusion of Primary Studies

8.2.2.1 Searching for the Evidence

Once the research question is formulated, the following stage is the beginning of the bibliographic **search for evidence**. The problem with searching in textbooks for an answer to this question is that by the time a book hits the shelves it is already out of date. It may even be that the bigger and more prestigious the book, the more out of date it is because of the enormous length of time it takes to get published. There are also real problems with relying primarily on a search engine such as Google Scholar. True, it has several benefits; it is free, easy to use, and brings up a very wide range of sources of information including the grey literature, but its potential weaknesses (less accurate than PubMed, less advanced options for fine-tuning searches, non-specificity of results, insufficient coverage of *Cochrane Systematic Reviews*, time delay in coverage) should make any user cautious about using it as the main source for answering clinical questions (Haroon and Phillips 2010).

The purpose of a search strategy is to identify all possibly relevant studies to the review. In brief, this stage consists of searching electronic databases, searching for missed data, and searching for the ‘grey’ (including unpublished or incompletely published) literature (Needleman 2002).

The solution when faced with a clinical question is to use one of a number of reliable (i.e. evidence based) medical resources.

Electronic searching: Online information systems available at the point of care can provide access to up-to-date evidence when a clinical question arises. They are potentially one of the most effective interventions to support evidence-based practice in a clinical setting (Westbrook et al. 2007).

Clinicians and educators currently utilize a variety of resources and interfaces to search the biomedical literature to answer clinical questions:

- PubMed (<http://www.pubmed.gov/>)
- EBSCOhost
- CINAHL

- Natural Standard (www.naturalstandard.com)
- Natural Medicines Comprehensive Database (http://www.naturalmedicines.com/member_home.asp)
- The US federal Agency for Healthcare Research and Quality (www.ahrq.gov)
- The US National Library of Medicine’s MedlinePlus (www.nlm.nih.gov/medlineplus/)
- The Database of Abstracts of Reviews of Effects (<http://www.crd.york.ac.uk>)
- The National Guideline Clearinghouse (<http://www.guideline.gov>)
- The American Dental Association Evidence-based Dentistry Web site (<http://ebd.ada.org>)
- Cochrane Library (<http://www.cochrane.org/>)
- DARE (<http://www.crd.york.ac.uk/crdweb/>)
- US national guidelines clearhouse (<http://www.guideline.gov/>)
- NICE (www.nice.org.uk)
- SIGN (www.sign.ac.uk)
- EMBASE (www.library.nhs.uk/)

Two important resources to be aware of are SUMSearch and TRIPdatabase. **SUMSearch** (available from www.sumsearch.uthscsa.edu) is a metasearch engine run by the University of Texas. It functions by searching several databases (MEDLINE, DARE, and NGC, among others) for papers relevant to a user defined search strategy (e.g. asthmatic AND steroids) and then refining and rerunning the search if too many or too few ‘hits’ occur. In addition to this, SUMSearch can give basic advice about the search strategy (e.g. dropping the ‘tic’ in *asthmatic* and the ‘s’ in *steroids* and replacing the dropped letters with ‘*’) and allows the user to focus the search for papers on therapeutics or diagnosis, etc. **TRIPdatabase** (Turning Research Into Practice) is a search engine (available from www.tripdatabase.com) which as well as searching for systematic reviews will also identify primary research studies on a given topic as well as synopses and ‘clinical questions’ (Haroon and Phillips 2010).

Limitations of electronic searching: Searching electronic databases has become increasingly accessible and may at first appear to be a perfect and reliable means of tracking down literature. However, even the most experienced searchers will miss some relevant literature (Bickley and

Glenny 2003). The sensitivity (proportion of the total number of known randomized clinical trials identified by the search) of searching for ophthalmology randomized clinical trials published in 1988 was 82%, when the gold standard was for any journal, 87% for any journal indexed in MEDLINE, and 88% for selected journals indexed in MEDLINE. Weighted means for sensitivity across all studies were 51%, 77%, and 63%, respectively. The weighted mean for precision (proportion of publications retrieved by MEDLINE that were actually randomized clinical trials) was 8% (median 32.5%) (Dickersin et al. 1994). Specifically, in oral health, Bickley and Harrison (2003) searched four leading orthodontic journals on MEDLINE using the indexing terms randomized controlled trial or controlled clinical trial in the Publication Type field and compared the results to handsearching these four journals. The MEDLINE search identified 143 citations and handsearching of the journals identified 266 citations. The MEDLINE search identified 105 of the 266 (39.5%) citations that had been found by handsearching journals together with an additional 38 records. The 38 unmatched MEDLINE citations (12.5%) were examined, and 32 (84.2%) of these were found not to be *in vitro* controlled clinical trials. These included 17 *in vitro*, 5 retrospective, and 4 matched control studies, cross-sectional, and non-clinical studies. The remaining six citations (15.8%) were found to be controlled clinical trials, and these had been missed by handsearchers (Bickley and Harrison 2003).

The reasons for missed articles on searching MEDLINE are manifold. Essentially, for the researcher, electronic searches chiefly rely on two things: (a) the controlled vocabulary (in MEDLINE MeSH) terms assigned to the article by professional indexers and (b) descriptors (text words) used by author/s in the title and abstract. Lack of detail in the title and abstract of the paper will influence the results of a search, and as by no means all articles have abstracts, this limits the search potential even further. Another source of trials that slip through the electronic search net is conference abstracts that are published in journals but are not indexed in

the main bibliographic databases such as MEDLINE. Another limitation of electronic searching is that not all journals are indexed in MEDLINE and, in particular, non-English language references are under-represented (Bickley and Glenny 2003).

Handsearching to identify trials: Given the limitations of electronic searching, where comprehensive searching is paramount, then electronic searching must be supplemented by handsearching of key journals. Handsearching involves searching a journal page by page to identify all reports of controlled clinical trials, whether as full papers, abstracts, or correspondence (Bickley and Glenny 2003), in identifying trials reported as abstracts, letters, and those published in languages other than English, together with all reports published in journals not indexed in electronic databases (Hopewell et al. 2007). Checking reference lists to identify relevant studies for systematic reviews is frequently recommended by systematic review manuals and is often undertaken by review authors (Horsley et al. 2011).

The Cochrane Collaboration is dedicated to producing and disseminating systematic reviews of the effects of health care interventions to help people make well-informed decisions about health care. A comprehensive search for all relevant trials combining electronic searching with handsearching of key journals is essential to the validity of systematic reviews (Bickley and Glenny 2003).

8.2.2.2 The Grey Literature

Finally, most systematic reviews attempt to identify the so-called grey literature. This includes possibly relevant but not fully published studies (Needleman 2002).

8.2.2.3 Haynes's 5 S's

The concept of finding the low-hanging easily plucked fruit when searching for evidence has been conceptualized by Haynes (2006) (Fig. 8.2) to reflect the different sources of EBM available to clinicians and how they relate to each other hierarchically (Haroon and Phillips 2010).

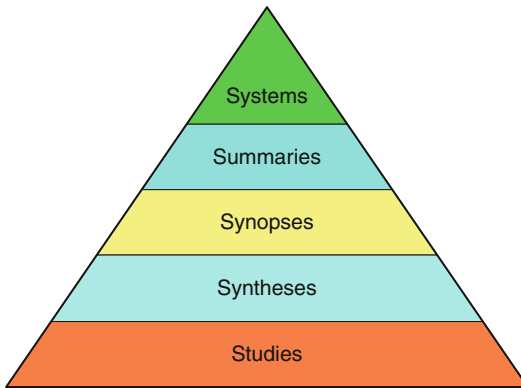


Fig. 8.2 Haynes's 5 S's. Illustrates the hierarchical relationship between different sources of evidence available to the clinician. Each tier up will generally represent a collation of the evidence below, and thus searching for an answer on the highest tier probably results in the most efficient search (Haroon and Phillips 2010. Reprinted with permission from BMJ Publishing Group Ltd.)

8.2.3 Study Selection Process

Study selection is a multistage process. Initially, the selection criteria are applied liberally to the citation generated from computed database searching. Unless they can be definitely excluded, the titles and abstracts identified as being potentially relevant from searches should be provisionally included for consideration on the basis of full-text articles (Manchikanti et al. 2009).

A typical process for selecting studies for inclusion in a reviews as follows (the process should be detailed in the protocol for the review) (Higgins and Deeks 2011):

1. Merge search results using reference management software and remove duplicate records of the same report.
2. Examine titles and abstracts to remove obviously irrelevant reports (authors should generally be over-inclusive at this stage).
3. Retrieve full text of the potentially relevant reports.
4. Link together multiple reports of the same study.
5. Examine full-text reports for compliance of studies with eligibility criteria.
6. Correspond with investigators, where appropriate, to clarify study eligibility (it may be

appropriate to request further information, such as missing results, at the same time).

7. Make final decisions on study inclusion and proceed to data collection.

Assessment of eligibility of studies, and extraction of data from study reports, should be done by at least two people, independently (Higgins and Deeks 2011). A final report of the review may also include a flow chart or a table detailing the studies included and excluded from the review, which are described in the text (Moher et al. 1999).

8.2.3.1 Study Quality Assessment

Quality is a multidimensional concept, which could relate to the design, conduct, and analysis of a trial, its clinical relevance, or quality of reporting. The validity of the findings generated by a study clearly is an important dimension of quality (Jüni et al. 2001). A convenient interpretation of quality is 'susceptibility to bias', although it is not uncommon for aspects of study conduct that are not directly associated with bias to be included in a quality assessment. For example, study size, whether or not a power calculation was performed, and ethical approval might be considered aspects of quality but are, in their own right, not potential causes of bias (Sanderson et al. 2007). It is necessary to take in consideration for the study quality, the degree to which a study employs measures to minimize biases (Manchikanti et al. 2009). Internal validity implies that the differences observed between groups of patients allocated to different interventions may, apart from random error, be attributed to the treatment under investigation. In contrast, external validity, or generalizability, is the extent to which the results of a study provide a correct basis for generalizations to other circumstances. In itself, there is no external validity. The term is only meaningful with regard to specified 'external' conditions, such as other patient populations or treatment regimens. Internal validity is a prerequisite for external validity: The results of a flawed trial are invalid, and the question of its external validity becomes redundant (Table 8.7) (Jüni et al. 2001).

8.2.3.2 Quality Assessment Instruments for Observational Studies

There are several methods in which one can assess the relationship between an intervention and an outcome. Randomized controlled trials

Table 8.7 Components of internal and external validity of controlled clinical trials (Jüni et al. 2001) (Reprinted with permission from BMJ Publishing Group Ltd.)

Internal validity – extent to which systematic error (bias) is minimized in clinical trials
• Selection bias: biased allocation to comparison groups • Performance bias: unequal provision of care apart from treatment under evaluation • Detection bias: biased assessment of outcome • Attrition bias: biased occurrence and handling of deviations from protocol and loss to follow up
External validity – extent to which results of trials provide a correct basis for generalization to other circumstances
• Patients: age, sex, severity of disease and risk factors, comorbidity • Treatment regimens: dosage, timing and route of administration, type of treatment within a class of treatments, concomitant treatments • Settings: level of care (primary to tertiary) and experience and specialization of care provider • Modalities of outcomes: type or definition of outcomes and duration of follow up

(RCTs) are considered as the gold standard for evaluating interventions. However, for many questions of clinical importance, RCTs would be impractical or unethical. Clinicians must rely on observational studies for the best available evidence when RCTs are unavailable (Lu 2009). **Table 8.8** presents a range of observational study designs; their characteristics, advantages, and methodological challenges are highlighted. Two major challenges of observational studies are selection bias and confounding. Selection bias is a systematic error due to design and execution errors in sampling, selection, or allocation methods. Selection bias can result in confounding. A factor can confound an association only if it differs between the intervention and comparison groups. For a variable to confound an association, it must be associated with both the intervention and outcome, and its relation to the outcome should be independent of its association with the intervention. Because observational studies often use data that were originally collected for other purposes, not all the relevant information may have been available for analysis. Thus, there are unknown potential confounders. Methods to improve the comparability of the intervention and control groups in observational study design include: (1) restriction – inclusion to the study is restricted to a certain category of a confounder

Table 8.8 Observational study designs (Lu 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Cohort studies follow one group that is exposed to an intervention of interest and another group that is non-exposed to determine the occurrence of the outcome (the relative risk). Cohort studies can examine multiple outcomes of a single exposure.
Case-control studies compare the proportion of cases with a specific exposure to the proportion of controls with the same exposure (the odds ratio). Case-control studies can examine multiple factors that may be associated with the presence or absence of the outcome.
Within-subject methods (case-only designs): <i>The self-controlled case-series</i> method assesses the association between a transient exposure and an outcome by estimating the relative incidence of specified events in a defined time period after the exposure. <i>Case-crossover</i> design estimates the odds of an outcome by comparing the probability of exposure between the exposure and control periods. <i>Case-time-control</i> design is case-cross over design with the addition of a traditional control group.
Cross-sectional studies are used to determine prevalence, that is, the number of cases in a population at a certain time and to examine the association between an exposure and an outcome.
Ecological studies focus on the comparison of groups. They can be used to identify associations by comparing aggregate data on risk factors and disease prevalence from different population groups.
Case reports provide anecdotal evidence by describing single cases. Description often includes the manifestations, clinical course, prognosis, how clinicians diagnosed and treated the condition and the clinical outcome.

Table 8.9 Major challenges of observational studies (Lu 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Selection bias*:	A systematic error in creating intervention groups, causing them to differ with respect to prognosis. The groups differ in measured or unmeasured baseline characteristics because of the way in which participants were selected for the study or assigned to their study groups.
Confounding*:	A situation in which the estimated intervention effect is biased because of some difference between the comparison groups apart from the planned interventions such as baseline characteristics, prognostic factors, or concomitant interventions. For a factor to be a confounder, it must differ between the comparison groups and predict the outcome of interest.
Strategies to reduce confounding	
<i>Design phase</i>	
Restriction: inclusion to the study is restricted to a certain category of a confounder (e.g., male).	
Matching of controls to cases to enhance equal representation of subjects with certain confounders among study groups.	
<i>Analytical phase</i>	
Stratification: the sample is divided into subgroups or strata on the basis of characteristics that are potentially confounding the analysis (e.g., age).	
Statistical adjustments	
Regression: estimates the association of each independent variable with the dependent variable (the outcome) after adjusting for the effects of other variables.	
Propensity score: a score that is the conditional probability of exposure to an intervention given a set of observed variables that may influence the likelihood of exposure.	
Instrumental variable: a pseudo-randomization method that divides patients according to levels of a covariate that is associated with the exposure but not associated with the outcome.	

*Definitions by the CONSORT statement from Rochon et al. [2005].

(e.g. male). However, strict inclusion criteria can limit generalizability of results to other segments of the population; and (2) matching of controls to cases (frequency matching or one-to-one matching) to enhance equal representation of subjects with certain confounders among study groups (Table 8.9) (Lu 2009).

Assessing quality and susceptibility to bias is essential when interpreting primary research and conducting systematic reviews and meta-analyses. It remains a dilemma that no gold standard exists with which to evaluate the quality (external and internal validity) of observational studies. A large variety of tools and checklists are available, but the large numbers of these probably reflect the lack of consensus between researchers as to which one is the best (Lang and Kleijnen 2010).

Agency for Healthcare Research and Quality (West et al. 2002) examined systematic approaches to assessing the strength of scientific evidence. Seventeen systems concerned observational studies. Of these, 4 were categorized as scales and 8 as checklists. The 8 domains used for assessing

these systems reflect general study design issues common to observational studies. Of these domains, 2 have empirical elements: comparability of subjects and funding or sponsorship. The remaining 7 domains – study question, study population, exposure/intervention, outcomes, statistical analysis, results, and discussion – come from best practices criteria. These domains are typically evaluated when critiquing an observational study.

Sanderson et al. (2007) reviewed the tools for assessing quality in observational epidemiological studies. A total of 86 tools were reviewed, comprising 41 simple checklists, 12 checklists with additional summary judgements, and 33 scales. The number of items ranged from 3 to 36 (mean 13.7). One-third of tools were designed for single use in a specific review and one-third for critical appraisal. Half of the tools provided development details, although most were proposed for future use in other contexts. Most tools included items for selection methods (92%), measurement of study variables (86%), design-specific sources of bias (86%), control of

confounding (78%), and use of statistics (78%); only 4% addressed conflict of interest. The distribution and weighting of domains across tools was variable and inconsistent.

Recently, Shamliyan et al. (2010) created a comprehensive evaluation of checklists and scales used to evaluate observational studies that examine incidence or prevalence and risk factors for diseases. Forty-six scales and 51 checklists were identified. Forty-seven of these tools were created for therapeutic studies, 48 for risk factors, and 5 for incidence studies. Forty-seven percent were modifications of previously published peer-reviewed appraisals, 18% were developed based on methodological standards, and 35% did not report development. Twenty-two percent reported reliability and 10% the validation procedure. Tools did not discriminate poor reporting versus methodological quality of studies or external versus internal validity; 35% categorize quality by the presence of predefined major flaws in design or by total score from the scale. Level of evidence was proposed in 22% of the tools by criteria of causality or internal validity of the studies. Evaluation required different degrees of subjectivity. Format, length, and content varied substantially across available checklists and scales.

Both reviews concluded that tools should be rigorously developed, evidence based, valid, reliable, and easy to use.

The Newcastle–Ottawa scale (NOS) is an ongoing collaboration between the Universities of Newcastle, Australia, and Ottawa, Canada (available at http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm). It was developed to assess the quality of nonrandomized studies with its design, content, and ease of use directed to the task of incorporating the quality assessments in the interpretation of meta-analytic results. A ‘star system’ has been developed in which a study is judged on three broad perspectives: the selection of the study groups, the comparability of the groups, and the ascertainment of either the exposure or outcome of interest for case–control or cohort studies, respectively. The goal of this project is to develop an instrument providing an easy and convenient tool for quality

assessment of nonrandomized studies to be used in a systematic review (Wells et al. 2001). The scale was adapted and used by Chambrone et al. (2010a) who systematically assessed the factors influencing tooth loss during long-term periodontal maintenance. The methodological quality of the observational studies was assessed using the Newcastle–Ottawa scale (NOS) (Wells et al. 2001). If all criteria of methodological quality were fulfilled within the domains, points (‘stars’) were assigned to the respective study. The NOS scale was adapted for the purpose of this review, and each study included could receive a maximum of 11 points. Studies with 9–11 points were arbitrarily considered as being of high, with 6–8 points of medium and with <6 points as being of low methodological quality (Table 8.10) (Chambrone et al. 2010a).

Recently, Kunnen et al. (2010) evaluated the possible relationship between periodontal disease and pre-eclampsia, a major pregnancy complication. A generalized inflammatory response plays an important role in the pathogenesis of pre-eclampsia. Because periodontal disease is a low-grade inflammatory state, periodontal disease might contribute to the pathogenesis of pre-eclampsia. Methodological quality was assessed using specific study-design-related checklists based on the quality assessment forms designed by the Dutch Cochrane Collaboration (<http://dcc.cochrane.org>). Case–control studies scoring five or more pluses (5 out of 8 items of the Cochrane checklist for case–control studies) were considered to be methodologically acceptable (den Hartog et al. 2008) and were included in this review (Table 8.11). For the quality assessment of cohort studies, the Dutch Cochrane checklist for cohort studies was used. Cohort studies scoring six or more pluses (6 out of 9 items of the Cochrane checklists for cohort studies) were considered methodologically acceptable (Table 8.12). Because there was no adequate checklist for cross-sectional studies at the Dutch Cochrane Collaboration, apart from a checklist for diagnostic tests, a quality assessment checklist had to be developed. This checklist was adapted from the quality checklist used for cohort studies and included items of checklists for cross-sectional

Table 8.10 Newcastle-Ottawa quality assessment scale (Adapted by Chambrone et al. 2010a) (Reprinted with permission from John Wiley & Sons, Inc.)

Selection	
1. Representativeness of the patients who experienced tooth loss during PM	(a) Truly representative of the average sample of patients who followed regular PM* (b) Somewhat representative of the average sample of patients who followed regular PM* (c) Selected group of patients (e.g. non-compliers) (d) No description of the derivation of the group
2. Selection of the patients who did not experience tooth loss during PM	(a) Drawn from the same community as the patients who experienced tooth loss* (b) Drawn from a different source (c) No description of the derivation of the patients who did not experience tooth loss
3. Ascertainment of tooth loss	(a) Secure record (e.g. patient records)* (b) Structured interview* (c) Written self report (d) No description
4. Demonstration that the number of teeth present after active periodontal therapy was reported in the study	(a) Yes* (b) No
Comparability	
1. Comparability of groups (patients) on the basis of the design or analysis	(a) All patients followed the same maintenance period* (b) All patients followed the same maintenance visits interval*
Outcome	
1. Assessment of tooth loss	(a) Prospective standard assessment (prospective studies)* (b) Record linkage (retrospective studies)* (c) Self report (d) No description
2. Was the number of teeth lost due to periodontal reasons after a minimum follow-up period of 5 years reported in the study?	(a) Yes * (b) No
3. Adequacy of follow up of patients	(a) Complete follow up – all subjects accounted for* (b) Subjects lost to follow up unlikely to introduce bias – small number lost $\geq 70\%$ follow up, or description provided of those lost)* (c) Follow up rate $< 70\%$ and no description of those lost (d) No statement (case series must be accounted in this category)
Statistics	
1. <i>Validity of statistical analysis</i>	(a) Valid* (b) Invalid (c) Unclear or not reported
2. <i>What was the unit of analysis?</i>	(a) Patient* (b) Site (i.e. tooth lost) (c) Unclear or not reported

Note: A study can be awarded a maximum of one star (*) for each numbered item within the Selection, Outcome and Statistics categories. A maximum of two stars can be given for Comparability

Table 8.11 Dutch Cochrane Collaboration quality assessment checklist for case-control studies* (Kunnen et al. 2010) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Canakci et al. (2004)	Oettinger-Barak et al. (2005)	Contreras et al. (2006)	Khader et al. (2006)	Kunnen et al. (2007)	Canakci et al. (2007)	Siqueira et al. (2008)	Lohsoonthorn et al. (2009)	Nabet et al. (2010)	Shetty et al. (2010)
Quality criteria										
Are the characteristics of the study group clearly defined?	+	–	+	+	+	+	+	+	+	+
Are the characteristics of the control group clearly defined?	+	–	+	+	+	+	+	+	+	+
Can selection bias sufficiently be excluded	+	+	–	+	–	–	+	+	+	–
Are the exposure and method of assessment clearly defined (periodontal disease)?	+	+	+	+	+	+	+	+	+	–
Are the outcome and method of assessment clearly defined (pre-eclampsia)?	+	+	+	+	+	±	+	+	±	+
Is the exposure assessed blinded to disease status?	?	–	–	+	–	?	+	+	±	–
Are the most important confounders identified and are these taken into consideration with respect to the study design and analysis?	+	–	–	+	+	+	+	+	+	+
Reported odds ratios or risk ratios with 95% CI?	+	–	+	+	+	+	+	+	+	+

Five or more plusses; methodologically acceptable

*Quality assessment according to the Dutch Cochrane Collaboration quality assessment checklist for case-control studies

Table 8.12 Dutch Cochrane Collaboration quality assessment checklist for cohort studies (Kunnen et al. 2010) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Boggess et al. (2003)	Meurman et al. (2006)	Srinivas et al. (2009)
Quality criteria^a			
Are the characteristics of the comparative study groups clearly defined?	+	–	+
Can selection bias sufficiently be excluded?	+	–	–
Are the exposure and method of assessment clearly defined (periodontal disease)?	+	+	+
Are the outcome and method of assessment clearly defined (pre-eclampsia)?	+	+	+
Is the outcome assessed blinded to the exposure status?	–	+	+
Is there a sufficient follow-up?	+	+	+
Can selective loss-to-follow-up sufficiently be excluded?	–	+	–
Are the most important confounders identified and are these taken into consideration with respect to the study design and analysis?	+	–	+
Reported odds ratios or risk ratios with 95% CI?	+	–	+

Six or more plusses: methodologically acceptable

^aQuality assessment according to the Dutch Cochrane Collaboration checklist for cohort studies

Table 8.13 Quality assessment checklist for cross-sectional studies, adapted from the Dutch Cochrane Collaboration quality checklist for cohort studies (Kunnen et al. 2010) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Castaldi et al. (2006)
Quality criteria^a	
Are the characteristics of the comparative study groups clearly defined?	–
Can selection bias sufficiently be excluded?	–
Are the exposure and method of assessment clearly defined (periodontal disease)?	+
Are the outcome and method of assessment clearly defined (pre-eclampsia)?	+
Is the outcome assessed blinded to the exposure status?	+
Are dropouts and reasons for dropout reported?	+
Can selective loss-to-follow-up sufficiently be excluded?	+
Are the most important confounders identified and are these taken into consideration with respect to the study design and analysis?	–
Reported odds ratios or risk ratios with 95% CI?	+

Five or more plusses: methodologically acceptable

^aQuality assessment adapted from the Dutch Cochrane Collaboration checklist for cohort studies

studies (diagnostic tests). Cross-sectional studies scoring five or more plusses (out of eight items of this checklist) were included in this review (Table 8.13).

8.2.3.3 Quality Assessment Instruments for RCTs

Randomized controlled trials (RCTs) are recognized as the gold standard to assess the efficacy

of health care interventions. However, it has been demonstrated that RCTs are not necessarily unbiased. Some studies have shown that inadequacy of certain important methodological items such as allocation concealment is associated with an exaggeration of the estimated treatment effect. Such bias in the conduct of RCTs may have serious consequences for patients' care and decision-making. This is why important concerns

about the poor quality of RCTs and the low reporting of important methodological details required to judge quality have been raised, especially from the mid-1990s. These concerns led to the publication of the Consolidated Standards of Reporting Trials (**CONSORT**) Statement in 1996 and updated in 2001 that aimed to improve the reporting of RCTs by providing guidelines (Dechartres et al. 2011).

Many tools have been proposed for assessing the quality of studies for use in the context of a systematic review and elsewhere. Most tools are **scales**, in which various components of quality are scored and combined to give a summary score, or **checklists**, in which specific questions are asked (Higgins and Altman 2011). These two types have been used interchangeably; however, they are actually quite distinct. Scales and checklists both include items measuring quality; however, with a scale, the responses to the individual items are summed to create an overall summary score representing trial quality (Olivo et al. 2008).

A high number of different scales or checklists others than those derived from the CONSORT Statement can be identified (Moher et al. 1995; Katrak et al. 2004; Olivo et al. 2008; Dechartres et al. 2011). In 1995, Moher et al. presented an annotated bibliography of scales and checklists developed to assess quality of randomized controlled trials (RCTs). Twenty-five scales and nine checklists have been developed to assess quality. The checklists are most useful in providing investigators with guidelines as to what information should be included in reporting RCTs. The scales give readers a quantitative index of the likelihood that the reported methodology and results are free of bias. There are several shortcomings with these scales related to little or no standard scale development techniques. Future scale development is likely to be most beneficial if questions common to all trials are assessed, if the scale is easy to use, and if it is developed with sufficient rigour.

Olivo et al. (2008) summarized the content, construction, areas of development, and psychometric properties of scales used to evaluate the quality of the randomized controlled trials

(RCTs) in health care research. The included studies accounted for 21 scales and their modifications including Jadad (Jadad et al. 1996), Maastricht (Verhagen et al. 1998a), Delphi list (Verhagen et al. 1998b), PEDro (Maher et al. 2003), Maastricht-Amsterdam List (MAL) (Furlan et al. 2002), van Tulder (van Tulder et al. 2003), Bizzini (Bizzini et al. 2003), Chalmers (Chalmers and Haynes 1994), Reisch (Reisch et al. 1989), Andrew (Andrew et al. 1990), Imperiale (Imperiale and McCullough 1990), Detsky (Detsky et al. 1992), Cho and Bero (Cho and Bero 1994), Balas (Balas et al. 1995), Sindhu (Sindhu et al. 1997), Downs and Black (Downs and Black 1998), Nguyen (Nguyen et al. 1999), Oxford Pain Validity Scale (OPVS) (Smith et al. 2000), Arrivé (Arrivé et al. 2000), CONSORT, and Yates scales (Yates et al. 2005). According to their Web of Science search, the Jadad scale was by far the most frequently cited and the most commonly used scale by the health care community. The authors pointed out that five of these scales (the Maastricht, Delphi list, MAL, van Tulder, and PEDro scales) are interrelated. The Maastricht scale was developed in the Department of Epidemiology of the University of Maastricht without using formal scale development techniques. The same group of authors decided to develop a methodological quality scale using formal techniques of scale development; thus, the Delphi list emerged. Since then, the Maastricht scale seldom has been used. The Cochrane Collaboration Back Group (CCBG) used the Delphi list as a basis for their analysis and added some items they found relevant for back pain. This list was then called the 'Maastricht-Amsterdam List' (MAL) (19 items) because of the cooperation between the two groups. Later, the CCBG updated the MAL and only considered 11 items. It is this list that is considered the 'van Tulder List' and has been used by the CCBG and many systematic reviews since 2003. In addition, the PEDro scale was derived from the Delphi list. Therefore, the Delphi list is the basis for most of the scales used in physical therapy, and its items are included in many of the scales. However, the scales derived from the Delphi list

(the MAL, van Tulder List, and PEDro) have added some items and did not consider further validation of these new scales (Olivo et al. 2008).

Many reviews specifically aimed to assess the quality of randomized controlled trials (RCTs). Dechartres et al. (2011) evaluated the quality of reporting in such reviews. Regarding assessment of quality, were noted whether the criteria for assessment were reported and defined. It was recorded if these criteria included the use of the CONSORT checklist (Altman et al. 2001), the Jadad scale (Jadad and McQuay 1996), the Delphi list (Verhagen et al. 1998b), the Chalmers scale (Chalmers and Haynes 1994), the PEDro scale (Maher et al. 2003), another quality scale or checklist, important methodology items, and if yes, all items that were used to assess the quality in the RCTs were systematically recorded. It was distinguished the quality of reporting (i.e. whether allocation concealment was reported) and the quality of methodology (i.e. whether allocation concealment was adequate). Whether a consensus or an interrater agreement between multiple assessors was reported was also noted. One hundred and seventy-seven reviews were found, published from 1987 to 2007, 58% of which were published after 2002. Of these, 131 (74%) focused on the quality of RCTs, 44 (25%) on quality of reporting, and 2 (1%) assessed both. The search strategy was well reported (92%). The criteria for assessment were reported in 97% of the reviews but were defined in only 38% (Table 8.14). Seventy-four different items and 26 different scales were identified (Table 8.15). Allocation sequence generation and concealment were reported in 41% and 40%, respectively, but their adequacy was assessed in 20% and 29%, respectively; scales were used in 40% and Consolidated Standards of Reporting Trials (CONSORT) checklist in 12%. It was concluded that the number of reviews assessing the quality of RCTs has dramatically increased over time but varied within medical areas. Although an improvement seems to appear in the reporting of the methodology, how quality is assessed still raises important issues. Heterogeneity of criteria and scales used (many of them being inadequate)

and the lack of definition may limit the relevance of these reviews. Dechartres et al. (2011) recommended that these reviews use well-defined criteria to assess quality of RCTs, based on the risk of bias tool developed by the Cochrane Collaboration.

The Collaboration's recommended tool for assessing risk of bias is neither a scale nor a checklist. It is a **domain-based evaluation**, in which critical assessments are made separately for different domains. The use of scales for assessing quality or risk of bias is explicitly discouraged in Cochrane reviews. While the approach offers appealing simplicity, it is not supported by empirical evidence. Calculating a summary score inevitably involves assigning 'weights' to different items in the scale, and it is difficult to justify the weights assigned. Furthermore, scales have been shown to be unreliable assessments of validity (Jüni et al. 1999), and they are less likely to be transparent to users of the review. It is preferable to use simple approaches for assessing validity that can be fully reported (i.e. how each trial was rated on each criterion) (Higgins and Altman 2011).

The recommended approach for assessing risk of bias in studies included in Cochrane reviews is a two-part tool, addressing the seven specific domains, namely, sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, and 'other issues'. The tool is summarized in Table 8.16. The tool was revised in late 2010 after an evaluation project. Each domain in the tool includes one or more specific entries in a 'risk of bias' table. Within each entry, the first part of the tool describes what was reported to have happened in the study, in sufficient detail to support a judgement about the risk of bias. The second part of the tool assigns a judgement relating to the risk of bias for that entry. This is achieved by assigning a judgement of 'low risk' of bias, 'high risk' of bias, or 'unclear risk' of bias (Higgins and Altman 2011).

The domains of sequence generation, allocation concealment, and selective outcome reporting

Table 8.14 Methodology characteristics of the methodological reviews as reported in the articles according to the year of publication (Dechartres et al. 2011) (Reprinted with permission from Elsevier)

Methodology of the study	Overall, <i>n</i> (%) <i>N</i> =177	Before 1997, <i>n</i> (%) <i>N</i> =26	During 1997–2001, <i>n</i> (%) <i>N</i> =49	After 2001, <i>n</i> (%) <i>N</i> =102
Reporting of the search strategy	162 (91)	22 (85)	44 (90)	96 (94)
Electronic search	129 (80)	20 (91)	34 (77)	75 (78)
MEDLINE	105 (81)	18 (90)	29 (85)	58 (77)
Cochrane Collaboration databases	43 (33)	0 (0)	11 (32)	32 (43)
EMBASE	27 (21)	1 (5)	8 (23)	18 (24)
Hand search	60 (37)	6 (27)	22 (50)	32 (33)
Search of bibliography	50 (31)	10 (45)	13 (29)	27 (28)
References of systematic reviews	38 (23)	5 (23)	14 (32)	19 (20)
Contact experts	21 (13)	2 (9)	9 (20)	10 (10)
Reporting of how relevant references were selected	67 (38)	9 (35)	17 (35)	41 (23)
Two or more people independently	30 (45)	1 (11)	8 (47)	21 (51)
One person alone	23 (34)	6 (67)	5 (29)	12 (29)
One person alone with a quality assurance control	7 (10)	1 (11)	3 (18)	3 (7)
Two or more people not independently	5 (7)	1 (11)	1 (6)	3 (7)
One person alone with a second person if difficulty	2 (3)	0 (0)	0 (0)	2 (5)
Reporting of inclusion criteria	156 (88)	21 (81)	40 (82)	95 (93)
Reporting of the reasons for exclusion	79 (45)	11 (42)	19 (39)	49 (48)
Flowchart of the selection of references	22 (12)	0 (0)	2 (4)	20 (20)
List of references included available	85 (48)	16 (61)	24 (49)	45 (44)
List of references excluded available	14 (8)	4 (15)	4 (8)	6 (6)
Reporting of the number of assessors per script	128 (72)	20 (69)	33 (67)	75 (73)
Two or more independent assessors per script	95 (74)	12 (60)	23 (70)	60 (80)
One assessor per script with interrater agreement on a random sample of articles with a second assessor	20 (16)	4 (20)	7 (21)	9 (12)
One assessor per script	13 (10)	4 (20)	3 (9)	6 (8)
Reporting of consensus between assessors	98 (55)	10 (38)	23 (47)	65 (64)
Reporting of interrater agreement	75 (42)	8 (31)	19 (39)	48 (47)
Reporting of piloting the form	24 (14)	1 (4)	4 (8)	19 (19)
Reporting of criteria used for quality assessment	172 (97)	26 (100)	46 (94)	100 (98)
Reporting of the definition for criteria used for quality assessment	67 (38)	9 (35)	11 (22)	47 (46)

should each be addressed in the tool by a single entry for each study. For blinding of participants and personnel, blinding of outcome assessment and for incomplete outcome data, two or more entries may be used because assessments generally need to be made separately for different out-

comes (or for the same outcome at different time points). Review authors should try to limit the number of entries used by grouping outcomes, for example, as ‘subjective’ or ‘objective’ outcomes for the purposes of assessing blinding of outcome assessment, or as ‘patient reported at

Table 8.15 Criteria used to assess the quality of trials in the methodological reviews according to the year of publication (Dechartres et al. 2011) (Reprinted with permission from Elsevier)

Quality assessment tool used	Overall, n (%) N=177	Before 1997, n (%) N=26	During 1997– 2001, n (%) N=49	After 2001, n (%) N=102
Specific items	119 (67)	19 (73)	32 (65)	68 (67)
Reporting of method to generate allocation sequence	73 (41)	13 (50)	19 (39)	41 (40)
Assessment of its adequacy	35 (20)	9 (35)	11 (22)	15 (15)
Reporting of allocation concealment	71 (40)	5 (19)	19 (39)	47 (46)
Assessment of its adequacy	52 (29)	5 (19)	16 (33)	31 (30)
Reporting of intent-to-treat analysis	51 (29)	6 (23)	9 (18)	36 (35)
Assessment of its adequacy	24 (14)	4 (15)	6 (12)	14 (14)
Reporting of details of the sample size calculation	45 (25)	8 (31)	13 (27)	24 (24)
Assessment of its adequacy	10 (6)	4 (15)	2 (4)	4 (4)
Reporting of blinding (without details)	43 (24)	5 (19)	10 (20)	28 (27)
Assessment of its adequacy	5 (3)	0 (0)	1 (2)	4 (4)
Reporting of blinding of patients	26 (15)	7 (27)	6 (12)	13 (13)
Assessment of its adequacy	4 (2)	1 (4)	1 (2)	2 (2)
Reporting of blinding of care providers	15 (8)	2 (8)	2 (4)	11 (11)
Assessment of its adequacy	2 (1)	0 (0)	0 (0)	2 (2)
Reporting of blinding of outcome assessors	37 (21)	5 (19)	12 (24)	20 (20)
Assessment of its adequacy	4 (2)	1 (4)	1 (2)	2 (2)
Reporting of justification of withdrawals and dropouts	43 (24)	9 (35)	12 (24)	22 (22)
Assessment of its adequacy	10 (6)	5 (19)	1 (2)	4 (4)
Reporting of statistical analysis	33 (19)	7 (27)	12 (24)	14 (14)
Assessment of its adequacy	33 (19)	7 (27)	12 (24)	14 (14)
Reporting of inclusion criteria	32 (18)	6 (23)	12 (24)	14 (14)
Assessment of its adequacy	0 (0)	0 (0)	0 (0)	0 (0)
Reporting of baseline comparison of groups	31 (17)	9 (35)	10 (20)	12 (12)
Assessment of its adequacy	13 (7)	5 (19)	3 (6)	5 (5)
Reporting of details of intervention	29 (16)	8 (31)	10 (20)	11 (11)
Assessment of its adequacy	11 (6)	4 (15)	3 (6)	4 (4)
Reporting of power	28 (16)	3 (11)	12 (24)	13 (13)
Assessment of its adequacy	2 (1)	1 (4)	0 (0)	1 (1)
Reporting of the definition for the primary outcome	24 (14)	4 (15)	7 (14)	13 (13)
Assessment of its adequacy	8 (4)	1 (4)	4 (8)	3 (3)
Reporting of the proportion of loss to follow-up	15 (8)	3 (11)	5 (10)	7 (7)
Assessment of its adequacy	6 (3)	2 (8)	1 (2)	3 (3)
Others	77 (43)	16 (61)	22 (45)	39 (38)
Scales or checklists (other than CONSORT)	71 (40)	10 (38)	18 (37)	43 (42)
Jadad Scale	32 (18)	0 (0)	8 (16)	24 (24)
Chalmers Scale	14 (8)	8 (31)	3 (6)	3 (3)
Pedro Scale	4 (2)	0 (0)	0 (0)	4 (4)
Delphi list	4 (2)	0 (0)	1 (2)	3 (3)
Others	27 (15)	3 (11)	9 (18)	15 (15)
Modified CONSORT Checklist	22 (12)	0 (0)	6 (12)	16 (16)

CONSORT Consolidated Standards of Reporting Trials

Table 8.16 The Cochrane Collaboration’s tool for assessing risk of bias (Higgins and Altman 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Domain	Support for judgement	Review authors’ judgement
Selection bias		
Random sequence generation	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Selection bias (biased allocation to interventions) due to inadequate generation of a randomized sequence.
Allocation concealment	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.
Performance bias		
Blinding of participants and personnel: Assessments should be made for each main outcome (or class of outcomes)	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.
Detection bias		
Blinding of outcome assessment: Assessments should be made for each main outcome (or class of outcomes)	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Detection bias due to knowledge of the allocated interventions by outcome assessors.
Attrition bias		
Incomplete outcome data: Assessments should be made for each main outcome (or class of outcomes)	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Attrition bias due to amount, nature or handling of incomplete outcome data.
Reporting bias		
Selective reporting	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Reporting bias due to selective outcome reporting.
Other bias		
Other sources of bias	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review’s protocol, responses should be provided for each question/entry.	Bias due to problems not covered elsewhere in the table.

6 months' or 'patient reported at 12 months' for incomplete outcome data. The same groupings of outcomes will be applied to every study in the review. The final domain ('other bias') can be assessed as a single entry for studies as a whole (the default in RevMan). However, it is strongly recommended that pre-specified entries be used to address specific other risks of bias. Such author-specified entries may be for studies as a whole or for individual (or grouped) outcomes within every study. **Table 8.17** provides criteria for making judgements about risk of bias from each of the seven domains in the tool. If insufficient detail is reported of what happened in the study, the judgement will usually be 'unclear risk' of bias. An 'unclear' judgement should also be made if what happened in the study is known, but the risk of bias is unknown, or if an entry is not relevant to the study at hand (particularly for assessing blinding and incomplete outcome data, when the outcome being assessed by the entry has not been measured in the study)(Higgins and Altman 2011).

A 'risk of bias' table is available in RevMan for inclusion in a Cochrane review as part of the 'characteristics of included studies' table. For each entry, the judgement ('low risk' of bias, 'high risk' of bias, or 'unclear risk' of bias) is followed by a text box for a description of the design, conduct, or observations that underlie the judgement. Two figures may be generated using RevMan for inclusion in a published review. First, a 'risk of bias graph' figure illustrates the proportion of studies with each of the judgements ('low risk', 'high risk', 'unclear risk' of bias) for each entry in the tool. Second, a 'risk of bias summary' figure presents all of the judgements in a cross-tabulation of study by. An alternative, and perhaps preferable, version of the first figure (the 'risk of bias graph') would be to restrict attention to studies in a particular important meta-analysis and to represent the proportion of information (rather than the proportion of studies) at low risk, unclear risk, and high risk of bias. The proportion of information may be measured by the sums of weights awarded to the studies in the meta-analysis. However, such plots cannot currently be produced within RevMan (Higgins and Altman 2011).

In periodontology, Antczak et al. (1986a) presented the criteria and scoring method 'for a system to evaluate the quality of randomized control trials (RCTs) in dental research based on published reports, an adaptation of the quality assessment system for use in evaluating clinical trials in dental research. The system assumes that two aspects of an RCT are necessary for the assessment of the quality of the research. They are (1) the study protocol (**Table 8.18**) and (2) data analysis and presentation. The system has been designed to quantify explicitly the quality of an RCT by assigning an arbitrarily defined set of weights to a list of items whose presence and correctness are assumed to reflect the quality of the research. If a study fulfils all the requirements, a score of 1.00 is obtained. Sixty points refer to the design and implementation of the study protocol, and the remaining 40 to the modalities of data analysis and presentation Antczak et al. (1986b).

The objective of the systematic review performed by Montenegro et al. (2002) was to assess the quality of RCTs in periodontology using evidence-based components. Following a detailed search, screening and quality assessments of RCTs were conducted in duplicate and independently. Quality of individual components was assessed. The checklist for quality appraisal included:

- Randomization: Adequate if generated by random number table (computer-generated or not), tossed coin, and shuffled cards. Unclear if the study refers to randomization but either does not adequately explain the method or no method was reported. Inadequate randomization methods include alternate assignment, hospital number, and odd/even birth date.
- Adequate allocation concealment methods included central randomization (e.g. by telephone to a pharmacy or trial office), pharmacy sequentially numbered/coded containers, and sequentially numbered, opaque envelopes. Concealment was unclear if the study referred to allocation concealment but either did not adequately explain the method or no method was reported. Inadequate concealment involved methods where randomization could not be concealed, such as

Table 8.17 Criteria for judging risk of bias in the “Risk of bias” assessment tool (Higgins and Altman 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Random sequence generation	
Selection bias (biased allocation to interventions) due to inadequate generation of a randomized sequence.	
Criteria for a judgement of ‘Low risk’ of bias.	The investigators describe a random component in the sequence generation process such as: • Referring to a random number table; • Using a computer random number generator; • Coin tossing; • Shuffling cards or envelopes; • Throwing dice; • Drawing of lots; • Minimization
Criteria for the judgement of ‘High risk’ of bias.	The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example: • Sequence generated by odd or even date of birth; • Sequence generated by some rule based on date (or day) of admission; • Sequence generated by some rule based on hospital or clinic record number. Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-random categorization of participants, for example: • Allocation by judgement of the clinician; • Allocation by preference of the participant; • Allocation based on the results of a laboratory test or a series of tests; • Allocation by availability of the intervention.
Criteria for the judgement of ‘Unclear risk’ of bias.	Insufficient information about the sequence generation process to permit judgement of ‘Low risk’ or ‘High risk’.
Allocation concealment	
Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.	
Criteria for a judgement of ‘Low risk’ of bias.	Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: • Central allocation (including telephone, web-based and pharmacy-controlled randomization); • Sequentially numbered drug containers of identical appearance; • Sequentially numbered, opaque, sealed envelopes.
Criteria for the judgement of ‘High risk’ of bias.	Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: • Using an open random allocation schedule (e.g. a list of random numbers); • Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or nonopaque or not sequentially numbered); • Alternation or rotation; • Date of birth; • Case record number; • Any other explicitly unconcealed procedure.
Criteria for the judgement of ‘Unclear risk’ of bias.	Insufficient information to permit judgement of ‘Low risk’ or ‘High risk’. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement – for example if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque and sealed.

Table 8.17 (continued)

Blinding of participants and personnel	
Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.	
Criteria for a judgement of 'Low risk' of bias.	Any one of the following: - No blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding; - Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'High risk' of bias.	Any one of the following: - No blinding or incomplete blinding, and the outcome is likely to be influenced by lack of blinding; - Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken, and the outcome is likely to be influenced by lack of blinding.
Criteria for the judgement of 'Unclear risk' of bias.	Any one of the following: - Insufficient information to permit judgement of 'Low risk' or 'High risk'; - The study did not address this outcome.
Blinding of outcome assessment	
Detection bias due to knowledge of the allocated interventions by outcome assessors.	
Criteria for a judgement of 'Low risk' of bias.	Any one of the following: - No blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding; - Blinding of outcome assessment ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'High risk' of bias.	Any one of the following: - No blinding of outcome assessment, and the outcome measurement is likely to be influenced by lack of blinding; - Blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome measurement is likely to be influenced by lack of blinding.
Criteria for the judgement of 'Unclear risk' of bias.	Any one of the following: - Insufficient information to permit judgement of 'Low risk' or 'High risk'; - The study did not address this outcome.
Incomplete outcome data	
Attrition bias due to amount, nature or handling of incomplete outcome data.	
Criteria for a judgement of 'Low risk' of bias.	Any one of the following: - No missing outcome data; - Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); - Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; - For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate; - For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size; - Missing data have been imputed using appropriate methods.

(continued)

Table 8.17 (continued)

Criteria for the judgement of 'High risk' of bias.	Any one of the following: - Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; - For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate; - For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size; 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomization; - Potentially inappropriate application of simple imputation.
Criteria for the judgement of 'Unclear risk' of bias.	Any one of the following: - Insufficient reporting of attrition/exclusions to permit judgement of 'Low risk' or 'High risk' (e.g. number randomized not stated, no reasons for missing data provided); - The study did not address this outcome.
<i>Selective reporting</i>	
Reporting bias due to selective outcome reporting.	
Criteria for a judgement of 'Low risk' of bias.	Any of the following: - The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way; - The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon).
Criteria for the judgement of 'High risk' of bias.	Any one of the following: - Not all of the study's pre-specified primary outcomes have been reported; - One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; - One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); - One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; - The study report fails to include results for a key outcome that would be expected to have been reported for such a study.
Criteria for the judgement of 'Unclear risk' of bias.	Insufficient information to permit judgement of 'Low risk' or 'High risk'. It is likely that the majority of studies will fall into this category.
Other bias	
Bias due to problems not covered elsewhere in the table.	
Criteria for a judgement of 'Low risk' of bias.	The study appears to be free of other sources of bias.
Criteria for the judgement of 'High risk' of bias.	There is at least one important risk of bias. For example, the study: - Had a potential source of bias related to the specific study design used; or - Has been claimed to have been fraudulent; or - Had some other problem.
Criteria for the judgement of 'Unclear risk' of bias.	There may be a risk of bias, but there is either: - Insufficient information to assess whether an important risk of bias exists; or - Insufficient rationale or evidence that an identified problem will introduce bias.

*Minimization may be implemented without a random element, and this is considered to be equivalent to being random

Table 8.18 Study protocol evaluation form (Antczak et al. 1986a) (Reprinted with permission from John Wiley & Sons, Inc.)

	Points		Points
2.1 Selection description		2.9 Patient blinded	
(a) Adequate	3	(a) Yes	8
(b) Fair	1.5	(b) No/unknown	0
(c) Inadequate	0	(c) Not applicable	
2.2 Rejection log		2.10 Observer blind to treatment	
(a) Yes	3	(a) Yes	8
(b) Partial	1.5	(b) Partial	4
(c) No/unknown	0	(c) No/unknown	0
		(d) Not applicable	
2.3 Therapeutic regimen definition		2.11 Observer blind to results	
(a) Adequate	3	(a) Yes	4
(b) Fair	1.5	(b) Partial	2
(c) Inadequate	0	(c) No/unknown	0
2.4 Placebo/control appearance		2.12 Testing randomization	
(a) Same	3	(a) Adequate	3
(b) Different	1.5	(b) Fair	1.5
(c) Unstated	0	(c) Inadequate	0
(d) Not applicable			
2.5 Placebo/control taste		2.13 Testing blinding	
(a) Same	3	(a) Adequate	3
(b) Different	1.5	(b) Fair	1.5
(c) Unstated	0	(c) Inadequate	0
(d) Not applicable		(d) Not applicable	
2.6 Follow-up schedule		2.14 Stopping rules	
(a) Adequate	3	(a) Adequate	3
(b) Fair	1.5	(b) Fair	1.5
(c) Inadequate	0	(c) Inadequate	0
2.7 Test of adherence to treatment		2.15 Prior estimate of sample size	
(a) Adequate	3	(a) Yes, reported	3
(b) Fair	1.5	(b) Partial	1.5
(c) Inadequate	0	(c) No/unknown	0
2.8 Randomization blind		2.16 Error measurement	
(a) Yes	10	(a) Adequate and blinded	3
(b) Partial	5	(b) Partial	1.5
(c) Fair	0	(c) No/unknown	0
(d) Inadequate	0		
Total possible points 63			

alternate assignment, hospital number, and odd/even birth date.

- Blinding of patient, caregiver, and examiner was considered separately. These were recorded as adequate, inadequate, unclear, or, for examiner blinding, not applicable if the study design precluded the possibility of blinding. Handling of withdrawals and drop-

outs was assessed by analysis of whether all patients who entered the trial were properly accounted for at the end. Where dropouts occurred, the use of analyses to allow for losses (such as intention to treat) was noted (Montenegro et al. 2002).

In a recent systematic review of randomized trials evaluating the effects of periodontal

treatment in the association between periodontitis with preterm birth (PB) and/or low birth weight (LBW), the methodological quality of the studies was assessed by focusing on the points described in the Cochrane Collaboration's tool for assessing the risk of bias [as referenced in Chap. 8.5 and detailed in Table 8.5c of the *Cochrane Handbook for Systematic Reviews of Interventions* 5.0.1 and detailed in Appendix S1 (Higgins and Green 2011)]: method of randomization and allocation criteria (i.e. adequate, inadequate, and unclear), blindness of examiners (yes, no, and unclear), and completeness of the follow-up. The risk of bias in the included studies was categorized as follows: (a) low risk of bias (plausible bias unlikely to drastically alter the results) – if all criteria were met; (b) unclear risk of bias (plausible bias that raises some doubt about the results) – if one or more criteria were partly met; and (c) high risk of bias (plausible bias that drastically weakens confidence in the results) – if one or more criteria were not met (Chambrone et al. 2011).

8.2.4 Data Extraction

The data collection form is a bridge between what is reported by the original investigators (e.g. in journal articles, abstracts, personal correspondence) and what is ultimately reported by the review authors. The data collection form serves several important functions: First, the form is linked directly to the review question and criteria for assessing eligibility of studies and provides a clear summary of these that can be applied to identified study reports. Second, the data collection form is the historical record of the multitude of decisions (and changes to decisions) that occurs throughout the review process. Third, the form is the source of data for inclusion in an analysis. Given the important functions of data collection forms, ample time and thought should be invested in their design. Because each review is different, data collection forms will vary across reviews. However, there are many similarities in the types of information that are important, and forms can be adapted from one review to the next (Higgins and Deeks 2011).

Items to consider in data collection or data extraction include (Higgins and Deeks 2011):

- Source
- Eligibility
- Methods (study design, total study duration, sequence generation, allocation sequence concealment, blinding, other concerns about bias)
- Participants
- Interventions
- Outcomes
- Results
- Miscellaneous

The decision between data collection using paper forms and data collection using electronic forms is largely down to review authors' preferences. Electronic systems have been developed that offer most of the advantages of both approaches (including the commercial SRS software) (see www.trialstat.com). If review authors plan to develop their own electronic forms using spreadsheet or database programs, it is recommended that (1) a paper form is designed first and piloted using more than one author and several study reports, (2) the data entry is structured in a logical manner with coding of responses as consistent and straightforward as possible, (3) compatibility of output with RevMan is checked, and (4) mechanisms are considered for recording, assessing, and correcting data entry errors (Higgins and Deeks 2011).

In a recent systematic review of randomized trials evaluating the effects of periodontal treatment in the association between periodontitis with preterm birth (PB) and/or low birth weight (LBW), data were collated into evidence tables and grouped according to the type of periodontal disease (Tables 8.19 and 8.20) (Chambrone et al. 2011).

Other checklists cover a broad range of types of research and can be found on the excellent Critical Appraisal Skills Programme website (<http://www.phru.nhs.uk/casp/appraisa.htm>) (Needleman et al. 2005).

The use of checklists with objective criteria helps to safeguard the quality of the quality appraisal process itself. The process of devising the checklist helps to ensure that all relevant quality issues are included in the assessment. Written,

Table 8.19 Patients' baseline characteristics reported/considered by each included study (Chambrone et al. 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Patients' baseline characteristics
López et al. (2002)	Antibiotic therapy, body weight, educations, marital status, maternal age, parity, periodontal status, previous pregnancy outcomes, primiparous, smoking, urinary infection and vaginosis
Jeffcoat et al. (2003)	Bacterial vaginosis, body mass index ethnicity/race, foetal fibronecting, marital status, maternal age, nulliparous, previous pregnancy history, periodontal status and smoking
López et al. (2005)	Education, marital status, maternal age, periodontal status, previous pregnancy history, primiparous and smoking
Michalowicz et al. (2006)	Dental and periodontal status, education, ethnicity/race, gestational age, maternal age, previous pregnancy history, and systemic conditions
Offenbacher et al. (2006)	Bacterial vaginosis, education, ethnicity/race, gestational age, insurance assistance, marital status, maternal age, maternal height, maternal weight pregnancy, previous periodontal status, pregnancy history, smoking, systemic conditions and use of medications
Sadatmansouri et al. (2006)	Antibiotic therapy, education, maternal age, maternal weight, periodontal status, previous pregnancy history, systemic conditions, urinary infection and vaginal infection
Tarannum and Faizuddin (2007)	Maternal age, periodontal status, previous pregnancy history and socioeconomic status
Newnham et al. (2009)	Alcohol intake, body mass index, drugs intake, education, ethnicity/race, maternal age, nulliparous, periodontal status, previous pregnancy history, smoking and systemic conditions
Offenbacher et al. (2009)	Education, ethnicity/race, insurance assistance, marital status, maternal age, maternal weight pregnancy, periodontal status, previous pregnancy history, smoking and systemic conditions
Radnai et al. (2009)	Education, Maternal age, occupation of mothers, place of residence, previous pregnancy history, and periodontal status
Jeffcoat et al. (2010)	Alcohol intake, ethnicity/race, marital status, maternal age, periodontal status, previous dental attendance and smoking
Macones et al. (2010)	Education, ethnicity/race, marital status, maternal age, periodontal status, previous pregnancy history, and severity of PD
Oliveira et al. (2010)	Education, ethnicity/race, marital status, maternal age, maternal weight, number of prenatal visits and periodontal conditions

piloted checklists reduce but can never completely eliminate individual subjectivity in decisions. Having a written list means that it is more likely that the quality assessors will be both consistent and repeatable (Needleman et al. 2005).

8.2.5 Synthesis and Summary of Study Results

Data synthesis in systematic reviews or meta-analyses can be achieved through a descriptive or non-quantitative synthesis, complemented by the use of formal statistical techniques. In addition to generating a summary of the effects of interven-

tions, it is an integral part of the data synthesis to investigate whether the effects are consistent across the included studies, and if not, to investigate the reasons for the differences (Manchikanti et al. 2009).

8.2.5.1 Descriptive or Non-quantitative Synthesis

The objective of a descriptive or non-quantitative review is to correlate and present the extracted data in a manner such that information about the characteristics (population, interventions, outcomes, and study quality) and results of the studies included in the review are summarized in a meaningful way. This is best done by tabulation,

Table 8.20 Characteristics of included studies – periodontitis studies evaluating the effects of periodontal treatment in the association between periodontitis with preterm birth (PB) and/or low birth weight (LBW) (Chambrone et al. (2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Participants	Methods	PD definition	Outcomes	OR, RR, or HR (95% CI)	Conclusions	Notes
López et al. (2002)	400 pregnant women were randomly assigned to two groups, but 351 patients, aged 18–35, between 9 and 21 weeks' gestation, with >18 teeth, completed the study	Medical records, interviews and ultrasound + full-mouth periodontal examination (oral hygiene status, BOP, PPD, CAL) by two calibrated periodontists, upon entering the study Test group – periodontal treatment (plaque control instructions, SRP and 0.12% chlorhexidine mouthrinse once a day until delivery) before 28 weeks of gestation. Control group – no treatment before delivery	Periodontal disease was defined as the presence of ≥4 teeth with PPD ≥4 mm and with CAL ≥3 mm at the same site	Test group = 163 PB: 2 LBW: 1 Control group = 188 PB: 12 LBW: 7	OR (unadjusted) – ITT Test versus control group PB = 5.0 (1.09–22.9)	This study concluded that “periodontal disease appears to be an independent risk factor for PB and LBW; and periodontal therapy significantly reduces the rates of PB and LBW in women with PD”	University-based (Chile) This study was supported by the Fondo de Investigaciones Científicas y Tecnológicas

Jeffcoat et al. (2003)	368 pregnant women with periodontitis were randomly assigned to three groups, but 366 patients, mean age 22.5, between 21 and 25 weeks' gestation, completed the study	Medical records and ultrasound + full-mouth periodontal examination by a hygienist supervised by a periodontist, between 21 and 25 weeks' gestation Test group 1 – SRP plus placebo capsule three times a day Test group 2 – SRP plus metronidazole 250 mg three times a day for 1 week. Control group – tooth cleaning and polishing plus placebo capsule three times a day	Periodontal disease was defined as the presence >3 sites with CAL ≥ 3 mm	Test group 1 = 123 PB < 37 weeks = 5 PB < 35 weeks = 1 Test group 2 = 120 PB < 37 weeks = 15 PB < 35 weeks = 4 Control group = 123 PB < 37 weeks = 11 PB < 35 weeks = 6	RR (adjusted) – ITT Test 1 versus control group PB < 37 weeks = 0.5 (0.2–1.3) PB < 35 weeks = 0.2 (0.02–1.4) Test 2 versus control group PB < 37 weeks = 1.4 (0.7–2.9) PB < 35 weeks = 0.7 (0.2–2.4)	This study concluded that “performing SRP in pregnant woman with periodontitis may reduce PB in this population”	University-based (USA)
Michalowicz et al. (2006)	823 pregnant women were randomly assigned to two groups, but 812, 16 years old, 17 weeks' gestation, 20 teeth, completed the study	Medical records, ultrasound and reports from patients + full-mouth periodontal assessments (PPD, CAL, BOP, PI, and calculus) at baseline Test group – periodontal treatment (plaque control instructions, SRP and monthly tooth polishing until delivery) Control group – plaque control instructions	PD was defined as ≥ 4 teeth with a PPD ≥ 4 mm and a CAL ≥ 2 mm, as well as BOP ≥ 35% of tooth sites	Treatment group = 407 PB < 37 weeks = 49 PB < 35 weeks = 22 PB < 32 weeks = 10 LBW < 2,500 g = 40/406 LBW < 1,500 g = 8/406 Control group = 405 PB < 37 weeks = 52 PB < 35 weeks = 26 PB < 32 weeks = 18 LBW < 2,500 g = 43/403 LBW < 1,500 g = 15/403	HR (adjusted) – ITT Test versus control group PB = 0.93 (0.63–1.37)	This study concluded that “treatment of periodontitis in pregnant woman did not significantly alter rates of PB, LBW, or foetal growth restriction”	University-based (USA) This study was supported by the National Institute of Dental and Craniofacial Research

(continued)

Table 8.20 (continued)

Study	Participants	Methods	PD definition	Outcomes	OR, RR, or HR (95% CI)	Conclusions	Notes
Offenbacher et al. (2006)	109 pregnant women were randomly assigned to two groups, but 74 completed baseline examinations and 67, ≥ 18 years old, < 22 weeks' gestation, with ≥ 20 teeth, completed the study	Medical records + full-mouth periodontal examination (GI, PI, PPD, Rec, BOP) by calibrated examiners (K -scores 0.94–1.0) + biologic samples collected Test group – periodontal treatment (plaque control instructions, SRP and tooth polishing) Control group – supragingival debridement	Periodontal disease was defined as the presence of ≥ 2 sites with PPD ≥ 5 mm plus CAL of 1–2 mm at ≥ 1 sites with PPDs ≥ 5 mm.	Test group = 35 PB = 9 Control group = 32 PB = 14	OR (adjusted) – ACA PB $\times$ intervention = 0.26 (0.08–0.85) PB $\times$ control = 3.8 PB $\times$ baseline extent of PD ≥ 5 mm = 1.22 (1.02–1.46) ^a	This study provided “further evidence supporting the potential benefits of periodontal treatment on pregnancy outcomes”	University-based (USA) The study was supported by Philips Oral Health care
Sadatmansouri et al. (2006)	30 pregnant women, aged 18–35, between the 13th and 20th weeks' gestation, were randomly assigned to two groups and completed the study	Medical evaluation not reported + full-mouth periodontal examination (PPD, CAL, BOP) Test group – periodontal treatment (plaque control instructions, SRP and 0.2% chlorhexidine mouthrinse once daily for a week) Control group – no treatment before delivery	Periodontal disease was defined as the presence of ≥ 4 teeth with ≥ 1 site with PPD ≥ 4 mm and CAL ≥ 3 mm.	Test group = 15 PB = 0 PB/LBW = 0 Control group = 15 PB = 3 PB/LBW = 4	NR-ITT	This study concluded that “periodontal therapy, phase I, results in a reduction in PLBW incidence rate”	University-based (Iran)

Tarannum and Faizuddin (2007)	200 non-smokers pregnant women were randomly assigned to two groups, but 180, aged 18–35 years, 9–21 weeks' gestation, ≥20 teeth, completed the study	Medical records and interviews + full-mouth periodontal examination (oral hygiene index, BOP, PPD, CAL) with a manual probe (UNC-15) Test group – periodontal treatment (plaque control instructions, SRP and 0.2% chlorhexidine mouthrinse twice daily until periodontal therapy was completed) before 28 weeks gestation Control group – plaque control instructions	Periodontal disease was defined as the presence of CAL ≥ 2 mm at ≥50% of examined sites	Test group = 91 PB = 45 LBW = 19 Control group = 89 PB = 68 LBW = 48	NR-ITT and ACA	This study concluded that “non-surgical periodontal therapy can reduce the risk for preterm births in mothers who are affected by periodontitis.”	University-based (India) Additional information was obtained after contact with author
Newnham et al. (2009)	1,087 pregnant women were randomly assigned to two groups, but 1,073, >16 years of age, 12–20 weeks' gestation, ≥20 teeth, completed the study	Medical and dental questionnaire + full-mouth periodontal examination (PPD, CAL, BOP, PI) with a automated controlled-force probe Test group – periodontal treatment (plaque control instructions and SRP) Control group – no treatment before delivery	Periodontal disease was defined as PPD ≥ 4 mm in depth at ≥12 probing sites in fully erupted teeth	Test group = 538 PB = 52 Control group = 535 PB = 50	OR (adjusted) – ITT Test versus control group PB = 1.05 (0.7–1.58)	The evidence provided by the present study “does not support the hypothesis that treatment of periodontal disease during pregnancy in this population prevents preterm birth”	Hospital-based (Australia) This study was supported by the National Health and Medical Research Council of Australia, the Women and Infants Research Foundation and Channel 7 Telethon

(continued)

Table 8.20 (continued)

Study	Participants	Methods	PD definition	Outcomes	OR, RR, or HR (95% CI)	Conclusions	Notes
Offenbacher et al. (2009)	1,806 pregnant women were randomly assigned to two groups, but 1,745, mean age = 25.3 years, <23 weeks' gestation, with ≥20 teeth, completed the study	Medical records and ultrasound before 16 weeks of gestation + full-mouth periodontal examination (PPD and CAL) by calibrated examiners (<i>K</i> -scores 0.75–1.0) Test group – periodontal treatment (plaque control instructions, SRP and tooth polishing) Control group – no treatment	Periodontal disease was defined as the presence of ≥3 sites with CAL ≥ 3 mm	Test group = 874 PB < 37 weeks = 91 PB < 35 weeks = 36 PB < 32 weeks = 20 LBW < 2,500 g = 72/872 LBW < 1,500 g = 15/872 Control group = 871 PB < 37 weeks = 73 PB < 35 weeks = 33 PB < 32 weeks = 14 LBW < 2,500 g = 71/866 LBW < 1,500 g = 13/866	OR-ITT PB = 1.21 (0.08–1.66)	This study concluded that “periodontal therapy as provided in this protocol did not reduce the incidence of preterm delivery at less than 37, 35, or 32 weeks of gestational age”	University-based (USA) The study was supported by a NIDCR and NCRR
Radnai et al. (2009)	89 non-smokers pregnant women were randomly assigned to two groups, but 83, mean age 29, completed the study	Medical and demographic questionnaire and ultrasound + full-mouth periodontal examinations (PPD, BOP) by a single calibrated examiner (ICC ≥ 0.94) Test group – periodontal treatment on third trimester (plaque control instructions, SRP and tooth polishing) Control group – no treatment before delivery	PD was defined as the presence of ≥1 site with PPD ≥ 4 mm and BOP for ≥50% of teeth	Test group = 41 PB = 10 LBW = 6 PB/LBW = 4 Control group = 42 PB = 22 LBW = 18 PB/LBW = 14	OR (unadjusted) – ACA Test versus control group PB = 3.4 (1.3–8.6) LBW = 4.3 (1.5–12.6) ^a PB/LBW = 4.6 (1.3–15.5) ^a	This study concluded that “periodontal treatment completed before the 35th week appeared to have a beneficial effect on birth weight and time of delivery”	Hospital-based (Hungary) The study was supported by the University of Szeged, Hungary

Jeffcoat et al. (2011)	322 pregnant women, 22.2% smokers, mean age 23.7 years, between 6 and 20 weeks' gestation were randomly assigned to two groups and completed the study	Medical records and ultrasound+ full-mouth periodontal examination (CAL) with a calibrated periodontal probe at baseline and 20 weeks later Test group – periodontal treatment (plaque control instructions and SRP) in the first trimester Control group – plaque control instructions	Periodontal disease was defined as ≥ 3 sites with CAL ≥ 4 mm.	Test group = 160 PB < 35 weeks = 73 Control group = 162 PB < 35 weeks = 85	OR (adjusted) – ITT Successful treatment $\times$ full-term birth = 6.01 (2.57–14.03) ^a	This study concluded that “pregnant women with PD should be offered conservative periodontal therapy, as it is safe, and, if successful, may reduce the incidence of spontaneous preterm birth”	University-based (USA) This study was supported by the Commonwealth of Pennsylvania and by Procter and Gamble Company Additional information was obtained after contact with author
Macones et al. (2010)	757 pregnant women were randomly assigned to two groups, but 720, between 6 and 20 weeks' gestation, completed the study	Medical records and ultrasound+ partial periodontal examinations (a maxillary and a mandibular quadrant for patients with >10 teeth) by trained nurse or dental hygienist Test group – periodontal treatment (plaque control instructions, SRP and tooth polishing) Control group – tooth polishing	Periodontal disease was defined as CAL ≥ 3 mm on ≥ 3 teeth. Moderate–severe periodontal disease was defined as CAL ≥ 5 mm on ≥ 3 teeth	Test group = 359 PB < 37 weeks = 58 ^b PB < 35 weeks = 31 ^b LBW < 2,500 g = 48 ^b /357 LBW < 1,500 g = 11 ^b /357 Control group = 361 PB < 37 weeks = 47 ^b PB < 35 weeks = 20 ^b LBW < 2,500 g = 35 ^b /359 LBW < 1,500 g = 6 ^b /359	RR (unadjusted) – ITT PB < 35 weeks = 1.61 (0.90–.88)/ Spontaneous PB < 35 weeks = 1.19 (0.62–2.28) PB < 37 weeks = 1.29 (0.85–.95)/ Spontaneous PB < 37 weeks = 1.03 (0.67–1.59)	This study concluded that “treating periodontal disease does not reduce the incidence of PB. There was a suggestion of an increase in the risk of indicated PB at <35 weeks of gestation in those subjects who received active treatment”	University-based (USA) This study was supported by the Pennsylvania Department of Health and by the National Center on Minority and Health Disparities Additional information was obtained after contact with author

(continued)

Table 8.20 (continued)

Study	Participants	Methods	PD definition	Outcomes	OR, RR, or HR (95% CI)	Conclusions	Notes
Oliveira et al. (2010)	246 non-smokers, pregnant women were randomly assigned to two groups, but 225, aged 18–35 years, 12–20 weeks' gestation, with ≥20 teeth, completed the study	Medical records, questionnaire and ultrasound+full-mouth periodontal examination (PPD, CAL and BOP) by a single calibrated examiner (K-scores ≥0.80) Test group – periodontal treatment (plaque control instructions, SRP when necessary and tooth polishing) Control group – supragingival debridement	Periodontal disease was defined as the presence of four or more teeth with one or more sites with PPD ≥4 mm and CAL ≥3 mm	Test group = 113 PB = 24 LBW = 23 PB/LBW = 29 Control group = 112 PB = 26 LBW = 31 PB/LBW = 31	RR (adjusted) – ACA Test versus control group PB = 0.91 (0.56–1.49) LBW = 0.73 (0.45–1.17) PB/LBW = 0.92 (0.60–1.43)	This study concluded that “nonsurgical periodontal treatment during the second semester of gestation did not significantly reduce the risk for the occurrence of PB, LBW, and PB/LBW”	University-based (Brazil) This study was supported by the Research Fund of Pontifical Catholic University of Minas Gerais

PPD probing depth, CAL clinical attachment level, RD recession depth, PI plaque index, BOP bleeding on probing, PD periodontal disease, PB preterm birth, LBW low birth weight, PB/LBW, preterm birth and low birth weight, OR odds ratio, RR relative risk, HR hazard ratio, ITT intention-to-treat analysis, ACA available cases analysis, CI confidence interval, NR not reported, NS non-significant

^aStatistically significant

^bData extracted from the tables (estimates)

which allows readers to look at the evidence, its methodological rigour, and the differences between the studies. The descriptive overview is an essential part of the data on which an understanding of the data, planning the quantitative data synthesis, and preventing errors in its interpretation are dependent (Manchikanti et al. 2009).

The process of carrying out the descriptive part of data synthesis should be explicit and rigorous. In general, the effectiveness of a health care intervention is dependent on a large number of factors, some known and others unknown, relating to who receives it, who delivers it and how, and in what context. The key elements in the descriptive approach to data synthesis may include multiple characteristics such as population; interventions; settings where the technology was applied; environmental, social, and cultural factors that may influence compliance; nature of the outcome measures used, their relative importance and robustness, the validity of the evidence; the sample sizes; and results of the studies included in the review (Manchikanti et al. 2009).

8.2.5.2 Quantitative Synthesis (Meta-Analysis)

Meta-analysis is not always possible when necessary data to perform meta-analysis cannot be obtained, and it may not be appropriate when the data are sparse or when the studies are too heterogeneous to be sensibly combined. The meta-analysis is performed to increase the power, to improve precision, and to answer the questions not posed by the individual studies, and to settle controversies arising from conflicting studies or to generate new hypothesis (Manchikanti et al. 2009; Egger et al. 1997).

Meta-analysis is a two-stage process. The first stage involves the calculation of a measure of treatment effect with its 95% confidence intervals (CI) for each individual study. The summary statistics that are usually used to measure treatment effect include odds ratios (OR), relative risks (RR), and risk differences. In the second stage of meta-analysis, an overall treatment effect is calculated as a weighted average of the individual summary statistics. The typical graph for displaying the results of a meta-analysis is called a 'forest plot' (Akobeng 2005).

Odds and Odds Ratio

The *odds* for a group is defined as the number of patients in the group who achieve the stated end-point divided by the number of patients who do not. For example, the odds of acne resolution during treatment with an antibiotic in a group of 10 patients may be 6–4 (6 with resolution of acne divided by 4 without=1.5); in a control group, the odds may be 3–7 (0.43). The *odds ratio*, as the name implies, is a ratio of two odds. It is simply defined as the ratio of the odds of the treatment group to the odds of the control group. In our example, the odds ratio of treatment to control group would be 3.5 (1.5 divided by 0.43) (Akobeng 2005).

Risk and Relative Risk

Risk, as opposed to odds, is calculated as the number of patients in the group who achieve the stated endpoint divided by the *total* number of patients in the group. Risk ratio or relative risk is a ratio of two 'risks'. In the example above, the risks would be 6 in 10 in the treatment group (6 divided by 10=0.6) and 3 in 10 in the control group (0.3), giving a risk ratio, or relative risk of 2 (0.6 divided by 0.3) (Akobeng 2005).

Interpretation of Odds Ratios and Relative Risk

An odds ratio or relative risk greater than 1 indicates increased likelihood of the stated outcome being achieved in the treatment group. If the odds ratio or relative risk is less than 1, there is a decreased likelihood in the treatment group. A ratio of 1 indicates no difference – that is, the outcome is just as likely to occur in the treatment group as it is in the control group. As in all estimates of treatment effect, odds ratios or relative risks reported in meta-analysis should be accompanied by confidence intervals (Akobeng 2005).

Readers should understand that the odds ratio will be close to the relative risk if the endpoint occurs relatively infrequently, say, in less than 20% (Akobeng 2005). If the outcome is more common, then the odds ratio will considerably overestimate the relative risk. The advantages and disadvantages of odds ratios versus relative risks in the reporting of the results of meta-analysis have been reviewed elsewhere (Akobeng 2005).

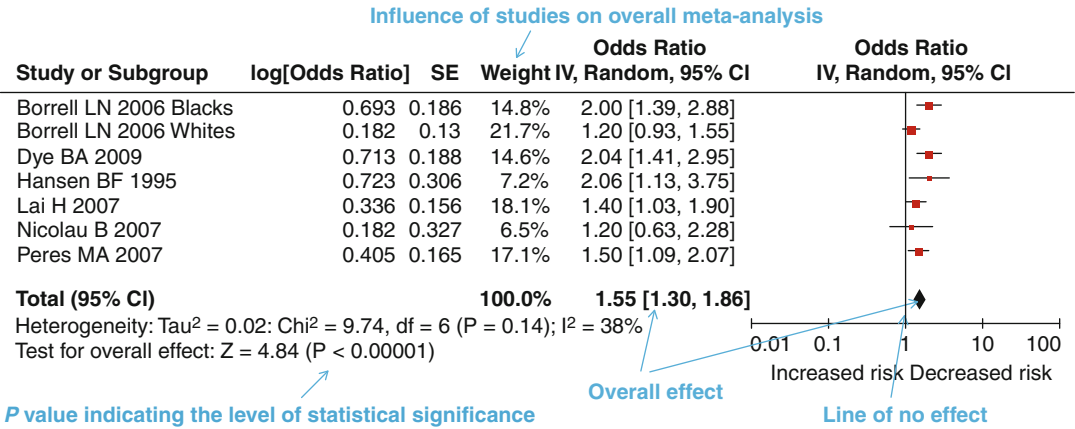


Fig. 8.3 Meta-analysis of binary outcome measures. Adjusted risk of periodontal diseases for individuals with low education. Legend: Higher level of education as reference. Weights for individual studies calculated with random effects models and inverse variance method. The centre of each red square is placed at the point estimate,

the area of the square is proportional to the sample size, and each horizontal line shows the 95% confidence interval for the estimate for each study. Pooled OR (95% CI): 1.55 (1.30–1.86). $P < 0.00001$ (Boillot 2011. doi:10.1371/journal.pone.0021508. Reprinted under the terms of the Creative Commons Attribution License)

Precision of the Treatment Effect: Confidence Intervals

Confidence intervals should accompany estimates of treatment effects. Ninety-five per cent confidence intervals are commonly reported, but other intervals such as 90% or 99% are also sometimes used. The 95% CI of an estimate (e.g. of odds ratios or relative risks) will be the range within which we are 95% certain that the true population treatment effect will lie. The width of a confidence interval indicates the precision of the estimate. The wider the interval, the less the precision. A very long interval makes us less sure about the accuracy of a study in predicting the true size of the effect. If the confidence interval for relative risk or odds ratio for an estimate includes 1, then we have been unable to demonstrate a statistically significant difference between the groups being compared; if it does not include 1, then we say that there is a statistically significant difference (Akobeng 2005).

Forest Plot

Meta-analysis results are commonly displayed graphically as ‘forest plots’. Figures 8.3 and 8.4 give examples of meta-analysis graphs. Figure 8.3 illustrates a graph with a binary outcome variable,

whereas Fig. 8.4 depicts a forest plot with a continuous outcome variable. The majority of meta-analyses combine data from randomized controlled trials (RCTs), which compare the outcomes between an intervention group and a control group. While outcomes for binary variables are expressed as ratios, continuous outcome measures are usually expressed as ‘weighted mean difference (WMD)’ in meta-analyses (Ried 2006).

The Heterogeneity Test

At the bottom of the graph on the left hand side, the number of interest is the I^2 value. I^2 was only recently developed and introduced as the preferable and more reliable test for heterogeneity. The principal advantage of I^2 is that it can be calculated and compared across meta-analyses of different sizes, of different types of study, and using different types of outcome data (Higgins et al. 2003). Other advantages of I^2 are:

- Focuses attention on the effect of any heterogeneity on the meta-analysis
- Interpretation is intuitive – the percentage of total variation across studies due to heterogeneity
- Can be accompanied by an uncertainty interval

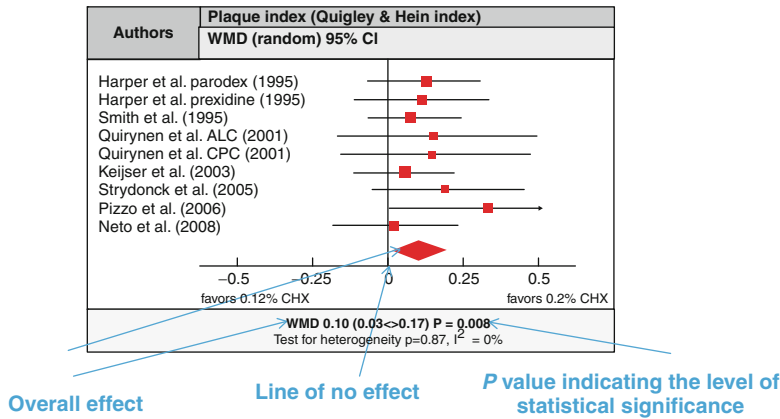


Fig. 8.4 Meta-analysis of continuous outcome measures. Meta-analyses Plaque Index comparing 0.12% versus 0.2% chlorhexidine. There was a significant difference in the effect on plaque between 0.12% CHX and 0.2% CHX. The WMD was 0.10 in favour of the 0.2% CHX mouthrinse

with a 95% confidence interval (CI: [0.03–0.17]). The test for overall effect showed a P -value of 0.008 (Berchier et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

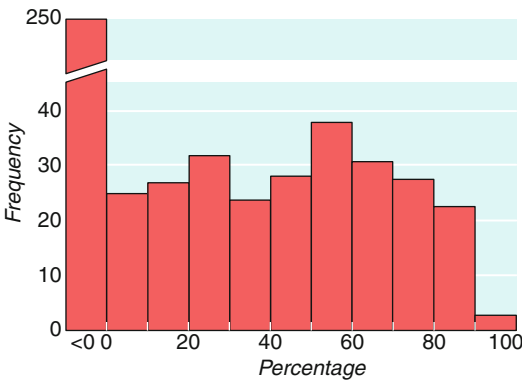


Fig. 8.5 Distribution of observed values of I^2 based on odds ratios from 509 meta-analyses of dichotomous outcomes in the *Cochrane Database of Systematic Reviews*. Data are from the first subgroup (if any) in the first meta-analysis (if any) in each review, if it involved a dichotomous outcome and at least two trials with events (Higgins et al. 2003. Reprinted with permission from BMJ Publishing Group Ltd.)

- Simple to calculate and can usually be derived from published meta-analyses
- Does not inherently depend on the number of studies in the meta-analysis
- May be interpreted similarly irrespective of the type of outcome data (e.g. dichotomous, quantitative, or time to event) and choice of effect measure (e.g. odds ratio or hazard ratio)
- Wide range of applications (Higgins et al. 2003)

Figure 8.5 shows the observed values of I^2 from 509 meta-analyses in the *Cochrane Database of Systematic Reviews*. Almost half of these meta-analyses (250) had no inconsistency ($I^2=0\%$). Among meta-analyses with some heterogeneity, the distribution of I^2 is roughly flat. A naive categorization of values for I^2 would not be appropriate for all circumstances, although we would tentatively assign adjectives of low, moderate, and high to I^2 values of 25%, 50%, and 75%. **Figure 8.5** also shows that about a quarter of meta-analyses have I^2 values over 50%. Quantification of heterogeneity is only one component of a wider investigation of variability across studies, the most important being diversity in clinical and methodological aspects. Meta-analysts must also consider the clinical implications of the observed degree of inconsistency across studies. For example, interpretation of a given degree of heterogeneity across several studies will differ according to whether the estimates show the same direction of effect (Higgins et al. 2003).

A useful visual guide to assess heterogeneity is to check the overlap of the CIs, that is, the horizontal lines or whiskers in the meta-analysis graph. Studies are regarded as homogeneous if

CIs of all studies overlap. Assessing inter- and intra-study variation or comparability of studies is important for the best choice of meta-analysis technique or model. Generally, one can choose between two models of meta-analysis, the ‘fixed’ and the ‘random effects’ models. If $I^2 \leq 25\%$, studies are regarded homogeneous, and the fixed effects model of meta-analysis can generally be used. If $I^2 \geq 75\%$, then heterogeneity is very high, and one should use a random effects model for meta-analysis (Ried 2006).

Fixed and Random Effects Models

Generally, a fixed effects model concentrates solely on the selected studies included in the meta-analysis, whereas a random effects model takes into account that there might be other studies unpublished, overlooked in the systematic literature search, or to be undertaken in the future which were not included in the meta-analysis at hand. Therefore, when the research question in the meta-analysis is whether treatment has produced an effect in the set of homogeneous studies analysed, then the fixed effects model is the appropriate one (Ried 2006).

Choosing the right model for analysis is particularly important if binary outcome variables are used, as fixed and random effects models give different results. In case of continuous variables, the results of meta-analyses using fixed or random models are often identical (Ried 2006).

8.2.6 Software Dedicated to Meta-analysis of Causal Studies

Bax et al. (2007) systematically assessed the differences in features, results, and usability of currently available meta-analysis programs. Six programs were included in their review: Comprehensive Meta-Analysis (CMA), MetAnalysis, MetaWin, MIX, RevMan, and WEasyMA. The programs differed substantially in features, ease of use, and price. Although most results from the programs were identical, some minor numerical inconsistencies were found. CMA and MIX scored highest on usability, and these programs also have the most complete set of analytical features (Bax et al. 2007). A summary of the meta-analysis software is presented in **Table 8.21**.

Table 8.21 Summary of meta-analysis software (Bax et al. 2007). doi: 10.1186/1471-2288-7-40 (Reprinted under the terms of the Creative Commons Attribution License)

	CMA	MetAnalysis	MetaWin	MIX	RevMan	WEasyMA
General						
URL	meta-analysis.com	–	metawinsoft.com	mix-for-meta-analysis.info	cc-ims.net/RevMan	weasyrna.com
Corporate single user price	~\$1295.00	~\$75.00	~\$150.00	Free	\$650	~\$490.00
Student single user price	~\$395.00	~\$75.00	~\$75.00	Free	Free	~\$280.00
Download/program size	30 MB	5 MB	9 MB	20 MB/50 MB	9 MB	3 MB
Compatibility	Windows	Windows	Windows	Windows	Windows	Windows
Last update	2006	2005	2002	2006	2005	2002
License	Single user	Single user	Single user	Open	Open	Single user
Input options						
Manual input	✓	✓	✓	✓	✓	✓
Copy and paste	✓		✓	✓	(✓)	
Text file import			✓	✓		

Table 8.21 (continued)

	CMA	MetAnalysis	MetaWin	MIX	RevMan	WEasyMA
File import (Excel, other software)			✓	✓		
Descriptive dichotomous, e.g. $n(\text{total})$, $n(y=1)$	✓	✓	✓	✓	✓	✓
Descriptive continuous, e.g. n , m , sd	✓		✓	✓	✓	
Comparative, e.g. θ , se , var	✓		✓	✓	✓	
Multi-format (mixed in one data set)	✓					
Single data input/selection	✓	✓	(✓)	✓	✓	✓
Maximum number of studies	Unlimited	Unlimited	Unlimited	100	Unlimited	Unlimited
Information sources						
Within-program HTML help			✓	(✓)	✓	(✓)
Printable manual	✓	✓	✓		✓	
Description of methods/calculations		(✓)	✓	(✓)	✓	
Additional information sources (PDFs/tutorials)	✓		✓	✓		
Up-to-date website	✓		✓	✓	✓	
Export options						
Copy output to clipboard	✓	✓	✓	✓	✓	✓
Export to office application(s)	✓					
Report creation	✓	✓		✓		
Setting copy file type (e.g. bmp, jpg or wmf)			✓	✓		✓

8.3 Reporting of Systematic Reviews and Meta-analysis of Observational Studies

The **MOOSE statement** (Stroup et al. 2000), a proposal for reporting of **Meta-analysis of Observational Studies in Epidemiology**, describes specifications of reporting including background, search strategy, methods, results, discussion, and conclusion. Use of the checklist should improve the usefulness of meta-analyses for authors, reviewers, editors, readers, and decision-makers. This document has not provided a flow diagram.

8.4 Reporting of Systematic Reviews and Meta-analysis of Randomized Controlled Trials

There is no guideline concerning the reporting of RCTs’ reviews. The **Quality of Reporting of Meta-analyses (QUOROM) Statement** (Moher et al. 1999) that has been updated recently and is now called **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)**. The PRISMA Statement consists of a 27-item checklist (**Table 8.22**) and a four-phase flow diagram (**Fig. 8.6**) (Liberati et al. 2009).

Some criteria of the PRISMA Statement may be of interest for RCTs’ reviews, especially those concerning the reporting of the search strategy, the selection of relevant references, and data extraction. These criteria are also included in the Overview Quality Assessment Questionnaire by Oxman and Guyatt (1991) and in the Assessment of Multiple Systematic Reviews (AMSTAR) tool (Shea et al. 2007), two other tools developed to assess the quality of systematic reviews. Search strategy, selection of references, and data extraction should also be clearly stated and explained in methodological reviews because readers need to assess the validity of the review and whether the conduct of the study is reproducible. Nevertheless, as goals of systematic and methodological reviews are noticeably different, some items of the AMSTAR tool (Shea et al. 2007) are not relevant for methodological reviews; for example, ‘were the methods used to combine the findings of studies appropriate?’ In systematic reviews, the aim is to identify and synthesize all studies concerning a particular question, whereas the aim of methodological reviews is to assess quality or reporting in a representative sample of studies. Therefore, being exhaustive is not as important as in a systematic review. Moreover, in the AMSTAR tool (Shea et al. 2007), there is no recommendation on how the scientific quality of studies should be assessed (Dechartres et al. 2011).

Table 8.22 Checklist of items to include when reporting a systematic review or meta-analysis(Liberati et al. 2009). doi: 10.1371/journal.pmed.1000100 (Reprinted under the terms of the Creative Commons Attribution License)

Section/topic	Item no	Checklist item	Reported on page no
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	
Abstract			
Structured summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	

Table 8.22 (continued)

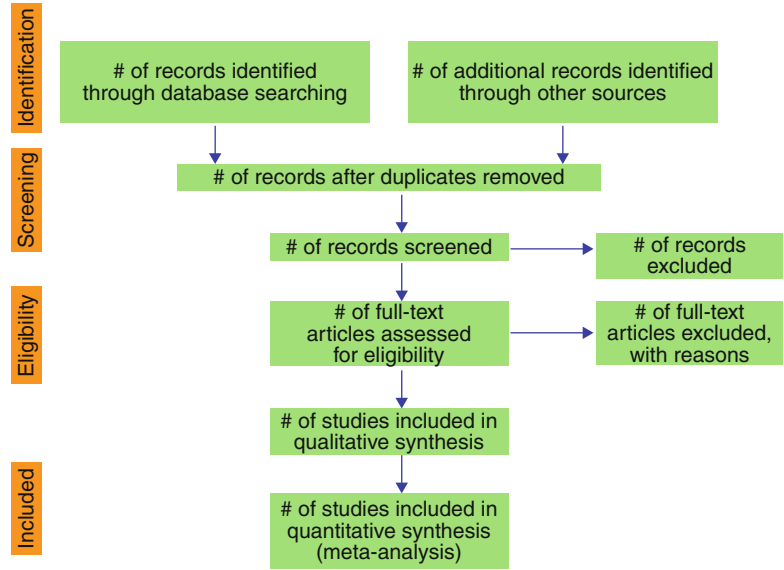
Section/topic	Item no	Checklist item	Reported on page no
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (such as web address), and, if available, provide registration information including registration number	
Eligibility criteria	6	Specify study characteristics (such as PICOS, length of follow-up) and report characteristics (such as years considered, language, publication status) used as criteria for eligibility, giving rationale	
Information sources	7	Describe all information sources (such as databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	
Study selection	9	State the process for selecting studies (that is, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	
Data collection process	10	Describe method of data extraction from reports (such as piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	
Data items	11	List and define all variables for which data were sought (such as PICOS, funding sources) and any assumptions and simplifications made	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	
Summary measures	13	State the principal summary measures (such as risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (such as I ²) for each meta-analysis	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (such as publication bias, selective reporting within studies)	
Additional analyses	16	Describe methods of additional analyses (such as sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	
Study characteristics	18	For each study, present characteristics for which data were extracted (such as study size, PICOS, follow-up period) and provide the citations	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot	

(continued)

Table 8.22 (continued)

Section/topic	Item no	Checklist item	Reported on page no
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15)	
Additional analysis	23	Give results of additional analyses, if done (such as sensitivity or subgroup analyses, meta-regression [see item 16])	
Discussion			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (such as health care providers, users, and policy makers)	
Limitations	25	Discuss limitations at study and outcome level (such as risk of bias), and at review level (such as incomplete retrieval of identified research, reporting bias)	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research	
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (such as supply of data) and role of funders for the systematic review	

Fig. 8.6 Flow of information through the different phases of a systematic review (Liberati et al. 2009. doi:10.1371/journal.pmed.1000100. Reprinted under the terms of the Creative Commons Attribution License)



8.5 Assessing the Quality of Systematic Reviews

Review articles are important sources of information to help guide decisions by clinicians, patients, and other decision-makers. Ideally, reviews

should include strategies to minimize bias and to maximize precision and be reported so explicitly that any interested reader would be able to replicate them. Compared with reviews published in paper-based journals, Cochrane reviews include elements that make them less prone to bias, such

Table 8.23 Criteria for assessing the scientific quality of research overviews (Oxman and Guyatt 1991) (Reprinted with permission from Elsevier)

Questions
1. Were the search methods reported?
2. Was the search comprehensive?
3. Were the inclusion criteria reported?
4. Was selection bias avoided?
5. Were the validity criteria reported?
6. Was validity assessed appropriately?
7. Were the methods used to combine studies reported?
8. Were the findings combined appropriately?
9. Were the conclusions supported by the reported data?
10. What was the overall scientific quality of the overview?

as the description of inclusion and exclusion criteria and formal assessment of trial quality. Also no Cochrane reviews had language restrictions (Jadad et al. 1998; Shea et al. 2002; Delaney et al. 2007). For rigorous systematic reviews, independent reviewers usually undertake quality appraisal in duplicate, and checklists are frequently employed for this purpose (Needleman 2005).

Assessment of systematic review quality can be performed using several quality assessment tools:

- The **Overview Quality Assessment Questionnaire (OQAQ)** (Oxman and Guyatt 1991; Oxman et al. 1991). The intended purpose of the criteria (Table 8.23) is to assess certain aspects of the scientific or methodological quality (i.e. adherence to scientific principles) of SRs published in the medical literature (to measure the extent to which a review is likely to be valid, i.e. the extent to which an overview guards against bias). The assessment of these aspects of scientific quality is primarily a measure of process rather than of outcome. It does not include a number of features that might be considered part of scientific quality. These include the importance of the question addressed, the degree of innovation in the approach, or the impact on future scientific and technical developments. The criteria were developed to assess threats to validity relative to each of these tasks (dimensions), specifically problem formulation, study identification, study selection, validation of studies, data extraction, data synthesis, and inference. After assessment of each of the nine component questions, a final

overall score was given, based on the answers to the previous questions on a scale of 1–7, 1 indicating extensive flaws, 7 indicating no flaws, a score ≤ 4 indicating major flaws, and a score ≥ 5 indicating that the study had only minimal flaws (Jadad and McQuay 1996; Lundh et al. 2009; Salmos et al. 2010; Chambrone et al. 2010b).

- The **checklist proposed by the Agency for Healthcare Research and Quality (AHRQ)** (West et al. 2002). The authors built on an earlier project concerning evaluating evidence for systematic reviews. They expanded this work by conducting a MEDLINE search (covering the years 1995 to mid-2000) for relevant articles published in English on either rating the quality of individual research studies or on grading a body of scientific evidence. Different instruments for critical appraising systematic reviews were reviewed, and 20 systems concerned with the appraisal of systematic reviews or meta-analysis were found, including 1 scale, 10 checklists, and 9 guidance documents, and identified 7 domains that they considered important to appraise: study question, search strategy, inclusion and exclusion criteria, data extraction, study quality, data synthesis and analysis, and funding or ownership (www.thecre.com/pdf/ahrq-system-strength.pdf).
- The **checklist proposed by Glenny et al. (2003)** is based on a combination of topics that have been demonstrated to influence the review's quality, as well as it covers all important issues that should be reported by high-standard SRs (i.e. from the identification of

Table 8.24 Quality assessment checklist for systematic reviews in dentistry used by Glenny et al. (2003). Per cent agreement range of κ values with asymptotic standard errors (not assuming the null hypothesis is true), assessing the inter-examiner agreement between clinicians and methodologists (Glenny et al. 2003) (Reprinted with permission from John Wiley & Sons, Inc.)

Question	Answers: possible categories	Percent agreement	κ values (asympt. SE)
A. Did review address a focused question?	Yes, no, can't tell	82	0.06 (0.13)
B. Did authors look for appropriate papers?	Yes, no, can't tell	72	0.38 (0.11)
C. Do you think authors attempted to identify all relevant studies?	Yes, no, can't tell	69	0.47 (0.09)
D. Search for published and unpublished literature	Yes, no, can't tell	71	0.46 (0.09)
E. Were all languages considered?	Yes, no, can't tell	88	0.81 (0.06)
F. Was any hand-searching carried out?	Yes, no, can't tell	71	0.40 (0.10)
G. Was it stated that the inclusion criteria were carried out by at least two reviewers?	Yes, no, can't tell	72	0.52 (0.09)
H. Did reviewers attempt to assess the quality of the included studies?	Yes, no	88	0.74 (0.09)
I. If so did they include this in the analysis?	Yes, no, can't tell, not applicable	57	0.36 (0.09)
J. Was it stated that the quality assessment was carried out by at least two reviewers?	Yes, no, not applicable	62	0.43 (0.08)
K. Are the results given in a narrative or pooled statistical analysis?	Narrative, pooled, not applicable	88	0.73 (0.08)
L. If the results have been combined was it reasonable to do so?	Yes, no, can't tell, not applicable	57	0.35 (0.09)
M. Are the results clearly displayed?	Yes, no, not applicable	80	0.60 (0.10)
N. Was an assessment of heterogeneity made and reasons for variation discussed?	Yes, no, not applicable	74	0.57 (0.09)
O. Were results of review interpreted appropriately?	Yes, no, can't tell, not applicable	55	0.25 (0.08)

clinical knowledge gap/focused question to results' interpretation) (Needleman et al. 2005). The aim of the study performed by Glenny et al. (2003) was to assess the quality of systematic reviews of effectiveness of interventions in dentistry. The Database of Abstracts of Reviews of Effectiveness and the Cochrane Database of Systematic Reviews were searched to identify systematic reviews examining the effectiveness of interventions for oral, dental, and craniofacial disorders and diseases. Sixty-five reviews were identified and assessed independently by two reviewers. The area most frequently evaluated within the reviews was pain relief/prevention (20/65, 31%)

followed by caries and oral medicine. One major weakness of the reviews was that the search strategies employed were not always adequate. Only 12 reviews (19%) demonstrated an attempt to identify all relevant studies. Other areas of weakness include the screening and quality assessment of primary studies, the pooling of data and examination of heterogeneity, and the interpretation of findings. This investigation shows that the quality of systematic reviews in dentistry could be improved. The checklist used in this study is reproduced here as an example. The items examined on the checklist included are listed in **Table 8.24**.

- The **AMSTAR (A Measurement Tool to Assess Systematic Reviews)** (Shea et al. 2007) was formed by combining (1) the enhanced Overview Quality Assessment Questionnaire (OQAQ), (2) a checklist created by Sacks, and (3) three additional items recently judged to be of methodological importance. The tool consists of 11 items and has good face and content validity for measuring the methodological quality of systematic reviews. AMSTAR was selected as the best instrument for appraising SRs by the Canadian Agency for Drugs and Technologies in Health due to its good reliability and convergent validity (**Table 8.25**) (Shea et al.

2007). The AMSTAR checklist has already been used to assess systematic reviews of dental studies (Bessa-Nogueira et al. 2008; Chambrone et al. 2010a; Faggion and Schmitter 2010; List and Axelsson 2010; Suebnukarn et al. 2010; Faggion et al. 2011). The results of the quality appraisal are used to assess the value of the evidence and to aid clinicians and reviewers in their efforts to place the evidence into context. This might be a part of the formal process of undertaking a systematic review or the informal act of reading and assessing recently published literature as part of everyday periodontal practice (Needleman 2005).

Table 8.25 AMSTAR is a measurement tool created to assess the methodological quality of systematic reviews (Shea et al. 2007), doi:10.1186/1471-2288-7-10 (Reprinted under the terms of the Creative Commons Attribution License)

1. Was an ‘a priori’ design provided? The research question and inclusion criteria should be established before the conduct of the review.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable
2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable
3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analysed e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable
7. Was the scientific quality of the included studies assessed and documented? ‘A priori’ methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.	☐ Yes ☐ No ☐ Can’t answer ☐ Not applicable

(continued)

Table 8.25 (continued)

8. Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigour and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
9. Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi-squared test for homogeneity, I^2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?).	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
10. Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
11. Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable

8.6 Pitfalls in Systematic Reviews

Although systematic reviews are considered the highest quality of evidence, there may be concerns about systematic reviews of which the reader should be aware (Farquhar and Vail 2006).

8.6.1 Failure to Adequately Assess Study Quality

The usefulness of findings from systematic reviews depends on the quality of the individual trials within the review. The concealment of allocation to the treatment group and the control group has been consistently shown to be the single most important factor in assessing the quality of RCTs. Trials in which the allocation is systematic, such as systems based on alternation, serial allocation, or days of the week, are considered to be inadequate in terms of the concealment of allocation. Blinding and concealment of allocation are frequently confused. Blinding is not always possible, especially in medical versus surgical trials, but even when possible it is often not performed (Farquhar and Vail 2006).

8.6.2 Failure to Take into Account Losses to Follow-Up

This is also known as **attrition bias**. Each study should report the patients who drop out and the reasons why, if known. It is possible that the reasons for dropping out could be due to adverse events, or poor outcomes, and the use of intention to treat analysis is usually recommended to address dropouts (Farquhar and Vail 2006).

8.6.3 Funding Bias

A systematic review has shown that published studies funded by commercial interests had four times the odds of favouring the experimental intervention, due to more careful selection of which studies to support, the narrow inclusion criteria that many industry trials apply, avoiding patients who have a poor prognosis, or patients with comorbidities (poor external validity), or selection of studies to submit for publication (Farquhar and Vail 2006).

8.6.4 Publication Bias

Publication bias can be seen to originate from three interrelated sources: researchers, journal editors, and research sponsors. Evidence suggests that researchers are less likely to ‘write up’ and submit a study if it is ‘negative’, and journal editors are less likely to publish this research if it is submitted. The common failure of authors to submit negative research may reflect a realistic appraisal of the fate of this research when submitted (probable rejection), rather than a wilful decision to suppress negative or uninteresting research. However, the sponsors of research may have a financial stake in the results of research into their own products – particularly in the field of pharmaceutical evaluation. Research which is sponsored by publicly funded bodies is likely to be published irrespective of its results, whereas commercially sponsored research is less readily published if the results are negative (Gilbody et al. 2000). However, there are a number of related biases that can also influence the availability and accessibility of research, even if it is published, such as the inclusion in systematic reviews and meta-analyses of tRCTs (truncated RCTs-tRCTs), failing to consider the possible overestimates of effect that may result from early stopping for benefit. Bassler et al. (2007) showed that of 96 systematic reviews that included at least one tRCT, 44 (46%) included >1 tRCT, 68 (71%) did not mention truncation at all, and 2 (2%) documented early stopping for benefit as a criterion for methodological quality. Of 47 meta-analyses in which authors reported, or we could calculate the contribution of the tRCTs to the pooled result, the tRCTs contributed more than 40% of the weight in 16/47 (34%).

Publication bias can be considered to have **three stages**: (1) **Prepublication bias** occurs in the performance of research, caused by ignorance, sloth, greed, or the double standard applied to clinical trials but not to clinical practice. (2) **Publication bias** refers to basing acceptance or rejection of a manuscript on whether it supports the treatment tested. Potentially biased reviewers are of equal concern. (3) **Postpublication bias** occurs in publishing interpretations, reviews, and

meta-analyses of published clinical trials (Chalmers et al. 1990).

Several analytical methods have been used to address the problem of publication bias.

8.6.4.1 Funnel Plot Analysis

A common approach is based on scatter plots of the treatment effect estimated by individual studies versus a measure of study size or precision (the ‘funnel plot’). In this graphical representation, larger and more precise studies are plotted at the top, near the combined effect size, while smaller and less precise studies will show a wider distribution below. If there is no publication bias, the studies would be expected to be symmetrically distributed on both sides of the combined effect size line. In case of publication bias, the funnel plot may be asymmetrical, since the absence of studies would distort the distribution on the scatter plot (Fig. 8.7) (Souza et al. 2007).

8.6.4.2 Begg’s Test: Rank Correlation Test

An adjusted **rank correlation test** was proposed by Begg and Mazumdar (1994) as a technique for identifying **publication bias** in a meta-analysis, and its **operating characteristics** are evaluated via simulations. The *test* statistic is a direct statistical analogue of the popular ‘funnel-graph’. The number of component studies in the meta-analysis, the nature of the selection mechanism, the range of variances of the effect size estimates, and the true underlying effect size are all observed to be influential in determining the power of the **test**. The **test** is fairly powerful for large meta-analyses with 75 component studies but has only moderate power for meta-analyses with 25 component studies. However, in many of the configurations in which there is low power, there is also relatively little **bias** in the summary effect size estimate. Nonetheless, the **test** must be interpreted with caution in small meta-analyses. In particular, **bias** cannot be ruled out if the **test** is not significant. The proposed technique has potential utility as an exploratory tool for meta-analysts, as a formal procedure to complement the funnel-graph (Begg and Mazumdar 1994).

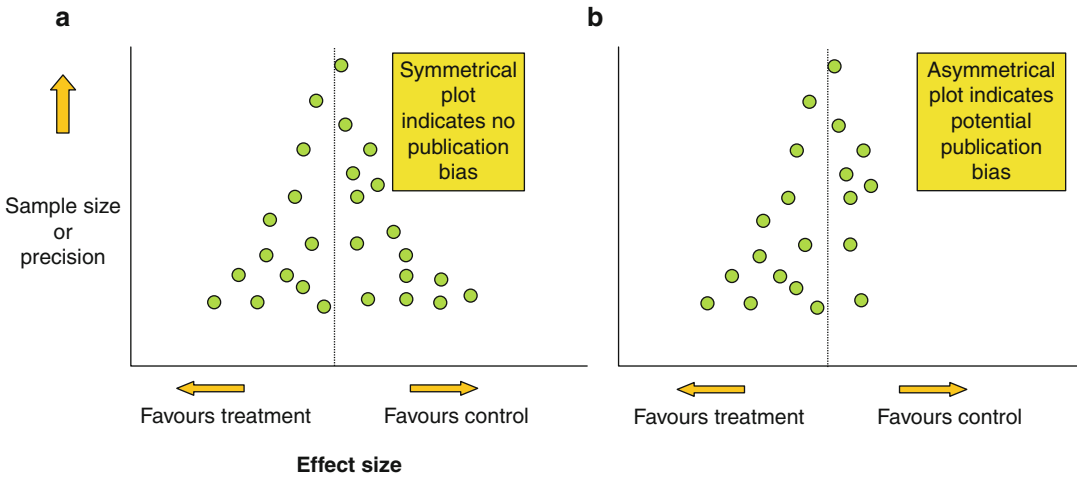


Fig. 8.7 Examples of symmetrical (a) and asymmetrical (b) funnel plots (Gilbody et al. 2000. Reprinted with permission from John Wiley & Sons, Inc.)

8.6.4.3 Egger's Linear Regression Approach

This technique is a formal test of funnel plot asymmetry and can be applied to datasets that report dichotomous outcomes or measures of association or risk, which are conventionally summarized using odds or risk ratios. The test uses three statistics: the log odds (or risk) ratio, its variance, and the standard normal deviate (SND). If there is no selection bias, the points from individual trials will scatter about a line that runs through the origin at standard normal deviate zero ($a=0$), with slope (b) indicating the size and direction of effect. If there is asymmetry, with smaller studies showing effects that differ systematically from larger studies, the regression line will not run through the origin. The intercept (a) provides a measure of asymmetry – the larger the deviation from zero, the more pronounced the asymmetry. Negative values ($-a$) will indicate that smaller studies show big protective effects (Gilbody et al. 2000).

8.6.4.4 Estimating the Number of Unpublished Studies

An alternative to the funnel plot is the 'file drawer problem' proposed originally by Rosenthal (1979). One of the descriptions of the overall problem is that 'researchers and statisticians have

long suspected that the studies published in the behavioural sciences are a biased sample of the studies that are actually carried out... The extreme view of this problem, the "file drawer problem," is that the journals are filled with the 5% of the studies that show Type I errors, while the file drawers back at the lab are filled with the 95% of the studies that show nonsignificant (e.g., $P>0.05$) results' (Rosenthal 1979).

The fundamental idea in coping with the file drawer problem is simply to calculate the 'fail-safe N', that is, the number of studies averaging null results that must be in the file drawers before the overall probability of a type I error is brought to any desired level of significance, say, $P=0.05$. This number of filed studies, or the tolerance for future null results, is then evaluated for whether such a tolerance level is small enough to threaten the overall conclusion drawn by the reviewer. If the overall level of significance of the research review will be brought down to the level of just significant by the addition of just a few more null results, the finding is not resistant to the file drawer threat (Rosenthal 1979).

For some areas of research, 100 or even 500 unpublished and unretrieved studies may be a plausible state of affairs, whereas for others, even 10 or 20 seem unlikely. Probably any rough and ready guide should be based partly on k so that as

more studies are known, it becomes more plausible that other studies in that area may be in those file drawers. Perhaps one could regard as resistant to the file drawer problem any combined results for which the tolerance level (X) reaches $5k + 10$. This seems a conservative but reasonable tolerance level; the $5k$ portion suggests that it is unlikely that the file drawers have more than five times as many studies as the reviewer, and the 10 sets the minimum number of studies that could be filed away at 15 (when $k=1$) (Rosenthal 1979).

The idea of the ‘fail-safe N ’ is to some degree sound, but there are two problems with it as an approach to the issue of publication bias. Firstly, it is a crude method for testing whether a significant result of meta-analysis can be made not significant by the addition of N studies that have an average null effect. This overemphasizes the importance of statistical significance, which is a disadvantage. Secondly, the method will always have a resultant effect in the same direction as the observed result of the meta-analysis. There are circumstances in which the unpublished studies have an average effect that is in the opposite direction to the observed meta-analysis, and when this happens, the fail-safe N is misleading.

8.6.4.5 TRIM and FILL Method

The ‘trim and fill’ method examines the existence of asymmetry in the funnel plot and is recommended as a tool for the assessment of the robustness of the results of meta-analyses (sensitivity analysis). The method consists of a rank-based data augmentation procedure that statistically estimates the number and location of missing studies. The main application of this method is to adjust for the possible effects of missing studies. If the conclusion of the meta-analysis remains unchanged following adjustment for the publication bias, the results can be considered reasonably robust, excluding publication bias (Souza et al. 2007).

Publication bias can be minimized by (1) insisting on high-quality research and thorough literature reviews, (2) eliminating the double standard concerning peer review and informed consent applied to clinical research and practice, (3)

publishing legitimate trials regardless of their results, (4) requiring peer reviewers to acknowledge conflicts of interest, (5) replacing ordinary review articles with meta-analyses, and (6) requiring the authors of reviews to acknowledge possible conflicts of interest (Chalmers et al. 1990). In addition, authors of systematic reviews should attempt to identify unpublished trials by contacting researchers and finally by using grey literature from conference abstracts. Updating reviews in a timely fashion is another way of avoiding publication bias (Farquhar and Vail 2006).

8.6.5 Considering the Outcomes of Interest

In many cases, research has focused on outcomes that may be of interest to the researcher but are of no value in clinical decision-making (Farquhar and Vail 2006). Endpoints are conditions or events that are associated with individual study subjects and that are used to assess treatment efficacy. Two types of endpoints can be distinguished: ‘true’ endpoints (reflect unequivocal evidence of tangible benefit to the patient) and ‘surrogate’ endpoints (usually a measure of disease process) (Hujoeel and DeRouen 1995).

True endpoints are outcomes that directly measure how a patient feels, functions, or survives. Tooth loss or painful periodontal abscesses are examples of events that can precisely be identified or realized by the patient’s mind, that is, that are tangible. True endpoints also include subjective oral health-related quality-of-life measurements or simple self-reported symptoms such as bleeding after brushing. Surrogate endpoints are intangible outcomes used as a substitute for a true endpoint (Hujoeel 2004). Surrogate endpoints save costs and time. Their disadvantages can be the uncertainty about their correlation to significant clinical outcomes, their inability to give an accurate picture of complications, especially long-term ones and, their inability, because they measure short-term changes, to address long-term effects such as implant longevity and, the variable ways in which they are measured (Paradis 2008). Two reasonable candidates for

surrogate endpoints for tooth loss are changes in probing attachment level and radiographic alveolar bone (height or mass) since substantial changes in either could essentially lead to tooth loss (Koch and Paquette 1997). Laboratory determinations of biological activity (e.g. microbiology, immunology) at periodontal sites are potential diagnostic and prognostic tests currently being considered (Koch and Paquette 1997; Loos and Tjoa 2005).

8.6.6 Analysis

The decision to pool study results by meta-analysis should be based on clinical rather than statistical criteria: Are the studies sufficiently similar in terms of the participants, the interventions, and the outcome assessment for an ‘average result’ to be interpretable? An important element in describing a review as ‘systematic’ is that, as far as possible, this decision is made before consideration of the trial results. In Cochrane reviews, the protocol is peer reviewed and published prior to assessment of the individual trials (Farquhar and Vail 2006).

8.6.7 Incorrect Use of Evidence Statements

There are two common errors in the interpretation of statistical ‘significance’ that occur in meta-analysis:

- The first is to confuse ‘significance’ with ‘importance’. A result is statistically ‘significant’ if the difference observed in the study sample is sufficiently convincing to signify a difference in the population of which the sample is representative. With a large enough sample size even a small and clinically unimportant sample difference will signify a population difference. The danger of assuming that a statistically significant difference establishes a clinically important difference is enhanced in meta-analysis, in which effective sample size is increased by pooling of multiple studies.

The second error is to misinterpret a statistically non-significant finding. The correct logic is to conclude that there is an absence of convincing evidence of an effect, not that there is evidence of absence of an effect. The importance of this misinterpretation should not be underestimated, as it can lead to the failure to further research or to take on innovative treatments (Farquhar and Vail 2006).

8.7 When Can a Meta-analysis Mislead?

- When a meta-analysis is done outside a systematic review
- When poor quality studies are included or when quality issues are ignored
- When small and inconclusive studies are included
- When inadequate attention is given to heterogeneity
- When reporting biases are a problem: publication bias, time lag bias, duplicate publication bias, language bias, outcome reporting bias (Egger et al. 2001)

8.8 Particular Types of Systematic Reviews and Meta-Analysis

Other types of systematic reviews and meta-analysis exist.

8.8.1 Realist Reviews

Realist reviews aim to determine how complex programs work in specific contexts and settings (Pawson et al. 2005).

8.8.2 Meta-Narrative Reviews

Meta-narrative reviews aim to explain complex bodies of evidence through mapping and comparing different over-arching storylines (Greenhalgh et al. 2005).

Table 8.26 Phases in meta-narrative review (Greenhalgh et al. 2005) (Reprinted with permission from Elsevier)

1. Planning phase (a) Assemble a multidisciplinary research team whose background encompasses the relevant research traditions (an initial scoping phase may be needed before the definitive research team is appointed). (b) Outline the initial research question in a broad, open-ended format. (c) Agree outputs with funder or client. (d) Set a series of regular face-to-face review meetings including planned input from external peers drawn from the intended audience for the review.	4. Appraisal phase Using appropriate critical appraisal techniques: (a) Evaluate each primary study for its validity and relevance to the review question; (b) Extract and collate the key results, grouping comparable studies together.
2. Search phase (a) Initial search led by intuition, informal networking and ‘browsing’, with a goal of mapping the diversity of perspectives and approaches. (b) Search for seminal conceptual papers in each research tradition by tracking references of references. Evaluate these by the generic criteria of scholarship, comprehensiveness and contribution to subsequent work within the tradition. (c) Search for empirical papers by electronic searching key databases, hand searching key journals and ‘snowballing’ (references of references or electronic citation tracking).	5. Synthesis phase (a) Identify all the key dimensions of the problem that have been researched; (b) Taking each dimension in turn, give a narrative account of the contribution (if any) made to it by each separate research tradition; (c) Treat conflicting findings as higher order data and explain in terms of contestation between the different paradigms from which the data were generated.
3. Mapping phase Identify (separately for each research tradition): (a) The key elements of the research paradigm (conceptual, theoretical, methodological and instrumental); (b) The key actors and events in the unfolding of the tradition (including main findings and how they came to be discovered); (c) The prevailing language and imagery used by scientists to ‘tell the story’ of their work.	6. Recommendations phase Through reflection, multidisciplinary dialogue and consultation with the intended users of the review: (a) Summarize the overall messages from the research literature along with other relevant evidence (budget, policymaking priorities, competing or aligning initiatives); (b) Distil and discuss recommendations for practice, policy and further research.

The phases of meta-narrative review are shown in **Table 8.26**. Themes are presented separate and sequential, but in reality, each phase overlapped with, and fed into, the next (Greenhalgh et al. 2005).

It was showed that while realist and meta-narrative reviews hold much promise for developing theory and informing policy in some of the health sector’s most pressing questions, misunderstandings and misapplications of these methods are common. A ‘RAMESES’ (Realist and Meta-review Evidence Synthesis: Evolving Standards) statement (comparable to CONSORT or PRISMA)

of publication standards for such reviews is under development (Greenhalgh et al. 2011).

8.8.3 Network Meta-Analyses

Network meta-analyses, also known as multiple treatments meta-analyses, can be used to analyse data from comparisons of many different treatments (Lumley 2002; Salanti et al. 2008). They use both direct and indirect comparisons and can be used to compare interventions that have not been directly compared.

Lumley (2002) has presented methods of estimating treatment differences between treatments that have not been directly compared in a randomized trial and, more importantly, methods of estimating the uncertainty in these differences. These methods require information from a large number of different treatment comparisons. The resulting estimates will be less reliable than those from direct randomized comparisons but may be useful in planning such comparisons or in cases where these direct comparisons are no longer ethically possible. To some extent, this lesser reliability will already be incorporated in our error estimates via the estimated incoherence (Lumley 2002).

Coherence of the network does not of itself guarantee that the conclusions are reliable and generalizable, but it provides a useful test for some important sources of bias and allows all the available information to be integrated (Lumley 2002).

Estimating incoherence is possible only when there are closed loops in the network of comparisons, and will be reliable to the extent that the trials in these closed loops are similar to other trials. This assumption can be weakened by modelling the incoherence more flexibly, though this would require different computational methods (Lumley 2002).

By incorporating all data presentations in a single analysis, we avoid the potential selection bias associated with conducting an analysis for a single statistic and the potential difficulties of interpretation, misleading results and loss of available treatment comparisons associated with conducting separate analyses for different summary statistics (Woods et al. 2010).

8.8.4 Prospective Meta-Analysis

Systematic reviews are by nature retrospective because the trials included are usually identified after the trials have been completed and the results reported. Knowledge of the results of individual randomized trials may introduce bias into a retrospective systematic review if the selection of the key components of the review question is

based on reports of one or more positive trials (Higgins and Green 2011).

A prospective meta-analysis (PMA) is a meta-analysis of studies (usually randomized trials) that were identified, evaluated, and determined to be eligible for the meta-analysis before the results of any of those studies became known. PMA can help to overcome some of the recognized problems of retrospective meta-analyses by:

- Enabling hypotheses to be specified a priori ignorant of the results of individual trials
- Enabling prospective application of study selection criteria
- Enabling a priori statements of intended analyses, including subgroup analyses, to be made before the results of individual trials are known

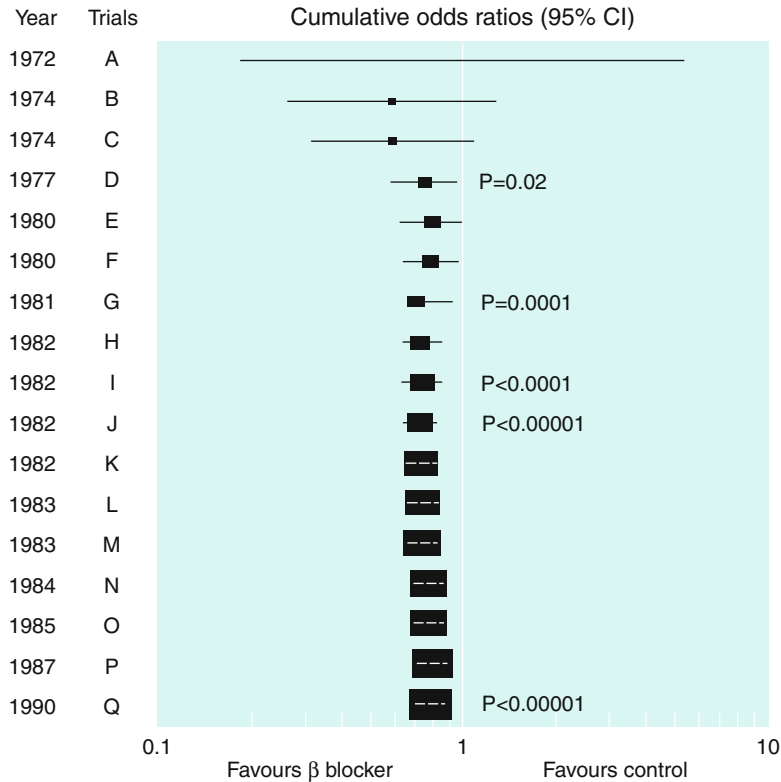
This avoids potential difficulties in interpretation related to the data-dependent emphasis on particular subgroups (Higgins and Green 2011).

Systematic reviews also depend on the ability of the review authors to obtain data on all randomized patients for the relevant outcomes, which can be difficult if full information is not reported in the trial publications. As most PMAs will collect and analyse individual patient data, they will be able to overcome this problem, with the additional advantage of being able to conduct time-to-event analyses if appropriate. Planned subgroup analyses based on patient-level factors can give misleading results if relying only on aggregate-level data, highlighting another advantage of individual patient data. PMA also provides a unique opportunity for trial design, data collection, and other clinical trial processes to be standardized across trials. For example, the investigators may agree to use the same instrument to measure a particular outcome and to measure the outcome at the same time points in each trial (Higgins and Green 2011).

8.8.5 Cumulative Meta-Analysis

Cumulative meta-analysis is defined as the repeated performance of meta-analysis whenever a new trial becomes available for inclusion. Such cumulative meta-analysis can retrospectively

Fig. 8.8 Cumulative meta-analysis of total mortality results from randomized controlled trials of oral β blockers after myocardial infarction. The size of the *square* reflects the amount of statistical information available at a given point in time (Egger and Smith 1997. Reprinted with permission from BMJ Publishing Group Ltd.)



identify the point in time when a treatment effect first reached conventional levels of significance (Egger and Smith 1997).

An example of this method is its application to the use of intravenous streptokinase as thrombolytic therapy for acute infarction (Lau et al. 1992). Thirty-three trials evaluating this therapy were performed between 1959 and 1988. It was found that a consistent, statistically significant reduction in total mortality (odds ratio, 0.74; 95% confidence interval, 0.59–0.92) was achieved in 1973, after only 8 trials involving 2,432 patients had been completed. The results of the 25 subsequent trials, which enrolled an additional 34,542 patients through 1988, had little or no effect on the odds ratio establishing efficacy but simply narrowed the 95% confidence interval. In particular, two very large trials, the Gruppo Italiano per lo Studio della Streptochinasi nell’Infarto Miocardico trial in 1986 (11,712 patients) and the Second International Study of Infarct Survival trial in 1988 (17,187 patients), did not modify the already established evidence of efficacy (Fig. 8.8).

Lau et al. (1992) used a similar approach to study the accumulating evidence of efficacy (or lack of efficacy) of 14 other therapies and preventive measures for myocardial infarction.

Another application of cumulative meta-analysis has been to correlate the accruing evidence with the recommendations made by experts in review articles and textbooks (Egger and Smith 1997).

The technique of performing cumulative meta-analyses is established in the medical literature. However, a MEDLINE search (October 2011) failed to find any reference to the use of the technique within any dental journal except of a demonstration exercise done by Moles et al. (2005). The starting point for the investigation was the systematic review of the effectiveness of adjunctive systemic antimicrobials for the treatment of periodontitis undertaken by the European Federation of Periodontology (Herrera et al. 2002). The focused question of that systematic review was, ‘In patients with chronic or aggressive periodontitis, what is the effect of systemic

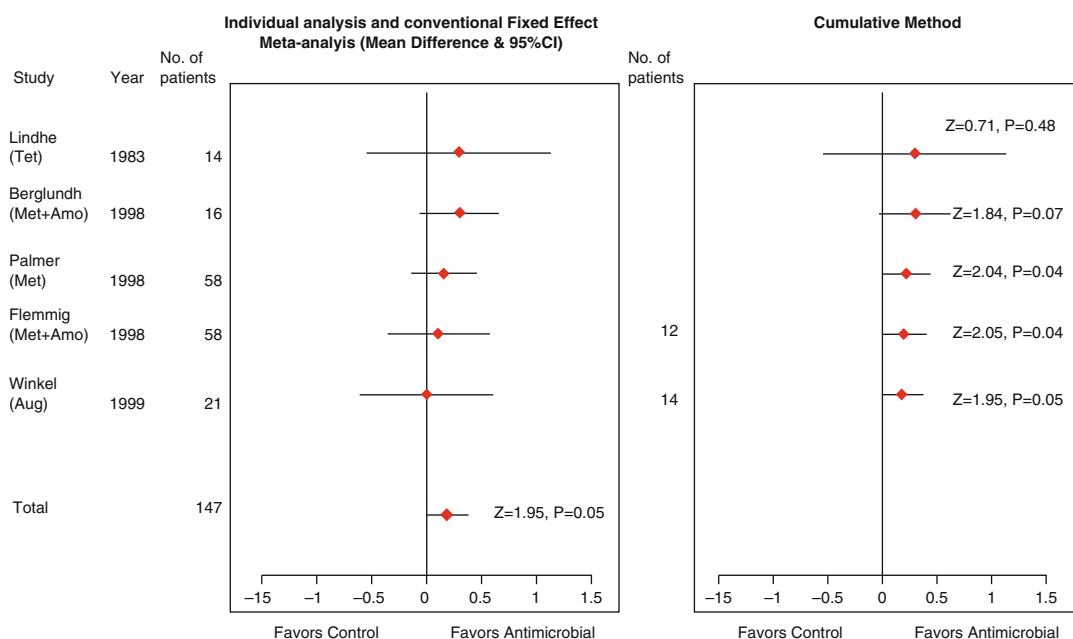


Fig. 8.9 Conventional and cumulative meta-analyses of five trials of systemic Antimicrobials as an adjunct to scaling and root planing for all/moderate periodontal pockets: difference in mean gain in clinical attachment level and 95% CI (SRP+antimicrobial vs. SRP alone) in mm. Key to figs.: Aug Augmentin®, Clin Clindamycin,

Dox Doxycycline, Met Metronidazole, Met+Amo Metronidazole and Amoxicillin in combination, Rod Rodogyl® (Metronidazole and Spiramycin in combination), Spi Spiramycin, Tet Tetracycline, Tet/Spi Tetracycline or Spiramycin (Moles et al. 2005. Reprinted with permission from Sage Publications)

antimicrobials as an adjunct to scaling and root planing (SRP+AB) compared with SRP alone, in terms of clinical outcomes?’ Conventional and cumulative meta-analyses were undertaken with the use of both fixed and random effects models, with the ‘MetaView’ software incorporated within the Cochrane ‘Review Manager’ program. This was achieved by use of an iterative approach, in which a series of separate meta-analyses was undertaken to include all the studies available at each information step. Thus, for each occasion that information became available, the new information was incorporated into the existing dataset, and the meta-analysis was repeated to update the estimates. This was undertaken separately for patient-level means of difference in CAL change in (a) all/moderate pockets at baseline (where moderate is defined as having a PPD 4–6 mm) and (b) deep pockets at baseline (PPD >6 mm). The overall effect size for adjunctive antimicrobials in all/moderate sites from the meta-analyses is a

mean difference of 0.18 mm (95% CI 0.00, 0.37; $P=0.05$) in favour of the adjunctive antimicrobial. The test for heterogeneity gave a chi-squared value of 1.00, df 4, $P=0.91$, showing no evidence of any difference in the effect between the different trials where the different antibiotic regimes were used. The cumulative meta-analysis indicated that statistical significance at the $P<0.05$ level was achieved marginally earlier in 1998 (Fig. 8.9) (Moles et al. 2005).

8.9 Improving and Updating Reviews

The definition of ‘to update’ means ‘to extend up to the present time’ or ‘to include the latest information’. Moher and Tsertsvadze (2006) defined the updating process as ‘a discrete event aiming to search for and identify new evidence to incorporate into a previously completed systematic

review'. The central and necessary element of an update is the effort to identify new evidence. We use 'new evidence' broadly – evidence that has not been included in the previously completed review. For example, use of the search strategy (e.g. MeSH terms, years searched) from the original review, but with an additional database (MEDLINE vs. MEDLINE and EMBASE) to identify new evidence, is regarded as an update. Alternatively, updating could be initiated after a specific period of time has elapsed since the completion of the original systematic review, which allows for the identification of new evidence that has emerged during this time (Moher and Tsertsvadze 2006).

An update can occur at any time after the completion of the original (or already updated) systematic review. For example, an update could be started immediately after the completion of a systematic review by using the same search strategy but including additional electronic databases. However, a more common practice is to update systematic reviews after a certain period has passed since the completion of the original version. Time could serve as a trigger or an indicator for the initiation of the updating process. A distinguishing feature of an updated systematic review from a new review is that during updating the originally formulated protocol (e.g. eligibility criteria, search strategy) is retained, and sometimes extended, to accommodate newly identified evidence (e.g. new treatment type, diagnostic method, outcome, different population) (Moher and Tsertsvadze 2006).

To date, there is little empirical evidence available to allow informed decisions about what is a reasonable and efficient approach to revisiting evidence in Cochrane reviews, although some guidelines do exist (Moher et al. 2007; Shojania et al. 2007a). The Cochrane Collaboration policy is that reviews should either be updated within 2 years or include a commentary to explain why this is not the case. In addition to the potential availability of new evidence, other developments may result in the need to revise a review. For example, within the clinical field, better tools or markers for characterizing subgroups may have been developed, new treatment regimens may be

available, or new outcome measures (or refined measurement methods of existing outcomes) may be in use. Furthermore, advances in the methods for conducting a Cochrane review may produce the need to revisit a review (Higgins and Green 2011).

The Cochrane Collaboration, the National Institutes of Health, and the Clinical Evidence update systematic reviews on a regular basis. Other systematic reviews, which account for about 80% of all published reviews, are not usually updated (Moher et al. 2007, 2008). Within 2 years of their publication, only 3% of systematic reviews published in peer-reviewed journals had been updated compared to 38% of Cochrane reviews (Moher et al. 2007, 2008; Shojania et al. 2007b; Garritty et al. 2010).

8.9.1 Strategies for Updating Systematic Reviews

Little research has been conducted on when and how to update SRs in contrast to other methodological areas of conducting SRs (e.g. publication bias, variance imputation). Moher et al. (2007) searched MEDLINE (1966 to December 2005), PsycINFO, the Cochrane Methodology Register, and the 2005 Cochrane Colloquium proceedings to identify records describing when and how to update SRs in health care. Four updating strategies, one technique, and two statistical methods were identified. Three strategies addressed steps for updating, and one strategy presented a model for assessing the need to update. One technique discussed the use of the 'entry date' field in bibliographic searching (Moher et al. 2007) (Table 8.27).

8.9.2 Statistical Methods for Updating Systematic Review

The statistical methods used for updating reviews identified by Moher et al. (2007) were cumulative meta-analysis and a test for detecting outdated meta-analyses with statistically non-significant results (Table 8.28).

Table 8.27 Strategies and techniques for updating systematic reviews – characteristics, strengths, and limitations (Moher et al. 2007) (Reprinted with permission from Elsevier)

Strategy (reference) country	Domains ^a	Strengths	Limitations
Steps in maintaining an updated review (Chalmers et al. 1993) UK	Search strategy Administrative issues	(a) Minimizes publication bias by obtaining grey literature and contacting authors for further information, clarifications	(a) Inefficient (b) Cumbersome to implement
Maintaining an updated review (Cochrane 2005, 2002)	International Search strategy Administrative issues Clinical Updating format	(a) Periodic updating ensures validity at some time (b) Timing is known (e.g., every 2 years)	(a) Inefficient (b) 2-year updating cycle may lead to outdatedness in rapidly developing fields, ^b or wasted resources in slowly developing fields
Assessment of need to update (Lutje et al. 2005) UK	Search strategy Administrative issues Clinical Public health	(a) Assesses the need to update (b) Prioritizes reviews requiring updating (c) Efficient (d) Evidence-based editorial consensus on whether or not to update (e) Algorithm of administrative actions	(a) Unclear how to determine whether a review is out of date (b) Unclear how to determine the importance of the topic in order to reach editorial consensus
Strategies for updating a review (Weller 1998) Australia	Clinical Public health Economic Updating format	(a) Applicable to systematic reviews, clinical practice guidelines, and health technology assessments	(a) General description of actions (b) Low practical utility
Searching using the “entry date” field (Bergerhoff et al. 2004) Germany	Search strategy	(a) Compensates for indexing lag by retrieving records indexed since the last search regardless of publication date	(a) It may not retrieve non-English records or those without abstracts

^aDomains: search strategy, statistical method/technique, clinical (expert opinion, long/short-term outcomes, intervention, pace of development in the field, nature of condition), public health importance (severity of condition and prevalence), economical (e.g., resource availability), updating format (preference for electronic vs. paper-based format; information steps), and administrative issues (e.g., those related to the implementation of updating process of systematic reviews)

^bMay not be the case if effects of complex interventions (e.g., educational, behavioural, other interventions at organizational level) are reviewed

Table 8.28 Statistical methods for updating systematic reviews- strengths and limitations (Moher et al. 2007) (Reprinted with permission from Elsevier)

Method (reference)	Strengths	Limitations
Identifying “null” meta-analyses that are ripe for updating (Barrowman et al. 2003)	(a) Relatively efficient (b) Easy to use/compute formula (c) Reduced type I error relative to conventional CMA (d) Test sensitivity/specificity easily modifiable	(a) Applicability limited to meta-analyses with statistically nonsignificant results (b) Assumes no secular trend in effect and that the variance of pooled estimate shrinks at a rate inversely proportional to the total number of participants in all studies (c) Test results sensitive to studies’ sizes
Conventional CMA (Lau et al. 1992, 1995; Baum et al. 1981; Berkey et al. 1996)	(a) Defines the earliest time at which an intervention can be shown to be efficacious, noninferior, or harmful (b) Monitors the effect size and direction over time (c) Timing for each update is known (d) Ascertains the contribution of individual studies to the cumulatively pooled effect estimate (e) Allows one to explore heterogeneity and perform sensitivity analysis (f) Provides up-to-date information (g) Useful in stopping ongoing trials or planning future trials	(a) Inefficient-if an update is conducted every time a new study becomes available (b) Inflated type I error due to multiple testing (c) Affected by publication bias
CMA using the cumulative slope as an indicator of stability (Mullen et al. 2001)	(a–g) of “Conventional CMA” (h) Explores the stability of the effect size and informs the need for updating	(a–c) of “Conventional CMA” (d) Judging extent of stability is arbitrary (e) The variance of the “cumulative slope” is invalid (f) The minimum size of a meta-analytic database for fitting a regression line whose slope would be a valid indicator of (in)stability of effect not specified
CMA using sequential monitoring boundaries (Pogue and Yusuf 1997)	(a–g) of “Conventional CMA” (h) Controls type I error by using sequential monitoring boundaries	(a and c) of “Conventional CMA” (d) Requires prior calculation of the OIS (e) Does not account for heterogeneity or bias among studies
Recursive CMA (Ioannidis and Lau 2001; Ioannidis et al. 1999)	(a–g) of “Conventional CMA” (h) Incorporates results from unpublished studies and follow-up or more detailed data for studies already included in the CMA (i) Documents the evolution of results as missing, updated, and new data are incorporated in information steps (j) Evaluates updated follow-up information, publication bias or lag, and heterogeneity (k) Treatment effect estimates are based on relatively accurate and complete data	(a and b) of “Conventional CMA” (c) Unpublished and updated information must be carefully studied and verified to minimize bias (d) Analysis of updated follow-up data may sometimes be inappropriate because many poststudy patients will cross over (e) More costly and resource consuming than conventional CMA

CMA cumulative meta-analysis, OIS optimal information size

References

- Akobeng AK. Understanding systematic reviews and meta-analysis. *Arch Dis Child*. 2005;90:845–8.
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663–94.
- Andrew E, Eide H, Fuglerud P, et al. Publications on clinical trials with X-ray contrast media: differences in quality between journals and decades. *Eur J Radiol*. 1990;10:92–7.
- Antczak AA, Tang J, Chalmers TC. Quality assessment of randomized control trials in dental research. II. Results: periodontal research. *J Periodontal Res*. 1986a;21:315–21.
- Antczak AA, Tang J, Chalmers TC. Quality assessment of randomized control trials in dental research. I. Methods. *J Periodontal Res*. 1986b;21:305–14.
- Armstrong EC. The well-built clinical question: the key to finding the best evidence efficiently. *Wis Med J*. 1999;98:25–8.
- Arrivé L, Renard R, Carrat F, et al. A scale of methodological quality for clinical studies of radiologic examinations. *Radiology*. 2000;217:69–74.
- Balas EA, Austin SM, Ewigman BG, et al. Methods of randomized controlled clinical trials in health services research. *Med Care*. 1995;33:687–99.
- Barrowman NJ, Fang M, Sampson M, Moher D. Identifying null meta-analyses that are ripe for updating. *BMC Med Res Methodol*. 2003;3:13.
- Bassler D, Ferreira-Gonzalez I, Briel M, Cook DJ, Devereaux PJ, Heels-Ansdell D, Kirpalani H, Meade MO, Montori VM, Rozenberg A, Schünemann HJ, Guyatt GH. Systematic reviewers neglect bias that results from trials stopped early for benefit. *J Clin Epidemiol*. 2007;60:869–73.
- Baum ML, Anish DS, Chalmers TC, Sacks HS, Smith Jr H, Fagerstrom RM. A survey of clinical trials of antibiotic prophylaxis in colon surgery: evidence against further use of no-treatment controls. *N Engl J Med*. 1981;305:795–9.
- Bax L, Yu LM, Ikeda N, Moons KG. A systematic comparison of software dedicated to meta-analysis of causal studies. *BMC Med Res Methodol*. 2007;7:40.
- Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50:1088–101.
- Berchier CE, Slot DE, Van der Weijden GA. The efficacy of 0.12% chlorhexidine mouthrinse compared with 0.2% on plaque accumulation and periodontal parameters: a systematic review. *J Clin Periodontol*. 2010;37:829–39.
- Bergerhoff K, Ebrahim S, Paletta G. Do we need to consider ‘in process citations’ for search strategies? In: 12th Cochrane colloquium in Ottawa, Ontario; 2004. p. 126.
- Bergus GR, Emerson M. Family medicine residents do not ask better-formulated clinical questions as they advance in their training. *Fam Med*. 2005;37:486–90.
- Berkey CS, Mosteller F, Lau J, Antman EM. Uncertainty of the time of first significance in random effects cumulative meta-analysis. *Control Clin Trials*. 1996;17:357–71.
- Bessa-Nogueira RV, Vasconcelos BC, Niederman R. The methodological quality of systematic reviews comparing temporomandibular joint disorder surgical and non-surgical treatment. *BMC Oral Health*. 2008;8:27.
- Bickley SR, Glenny AM. The cochrane oral health group trials register: electronic searching and beyond. *J Dent Educ*. 2003;67:925–30.
- Bickley SR, Harrison JE. How to.... find the evidence. *J Orthod*. 2003;30:72–8.
- Bizzini M, Childs JD, Piva SR, Delitto A. Systematic review of the quality of randomized controlled trials for patellofemoral pain syndrome. *J Orthop Sports Phys Ther*. 2003;33:4–20.
- Bogges KA, Lief S, Murtha AP, Moss K, Beck J, Offenbacher S. Maternal periodontal disease is associated with an increased risk for preeclampsia. *Obstet Gynecol*. 2003;101:227–31.
- Boillot A, El Halabi B, Batty GD, Rangé H, Czernichow S, Bouchard P. Education as a predictor of chronic periodontitis: a systematic review with meta-analysis population-based studies. *PLoS One*. 2011;6:e21508.
- Booth A. Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*. 2006;24:355–68.
- Bragge P. Asking good clinical research questions and choosing the right study design *Injury*. *Injury*. 2010;41S:S3–6.
- Canakci V, Canakci CF, Canakci H, Canakci E, Cicek Y, Ingec M, Ozgoz M, Demir T, Dilsiz A, Yagiz H. Periodontal disease as a risk factor for pre-eclampsia: a case control study. *Aust N Z J Obstet Gynaecol*. 2004;44:568–73.
- Canakci V, Canakci CF, Yildirim A, Ingec M, Eltas A, Erturk A. Periodontal disease increases the risk of severe pre-eclampsia among pregnant women. *J Clin Periodontol*. 2007;34:639–45.
- Castaldi JL, Bertin MS, Giménez F, Lede R. Periodontal disease: is it a risk factor for premature labor, low birth weight or preeclampsia? *Rev Panam Salud Publica*. 2006;19:253–8.
- Chalmers IHR, Haynes B. Reporting, updating, and correcting systematic reviews of the effects of health care. *BMJ*. 1994;302:862–5.
- Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA*. 1990;263:1392–5.
- Chalmers I, Enkin M, Keirse MJ. Preparing and updating systematic reviews of randomized controlled trials of health care. *Milbank Q*. 1993;71:411–37.
- Chambrone L, Chambrone D, Lima LA, Chambrone LA. Predictors of tooth loss during long-term periodontal maintenance: a systematic review of observational studies. *J Clin Periodontol*. 2010a;37:675–84.
- Chambrone L, Faggion Jr CM, Pannuti CM, Chambrone LA. Evidence-based periodontal plastic surgery: an assessment of quality of systematic reviews in the

- treatment of recession-type defects. *J Clin Periodontol*. 2010b;37:1110–8.
- Chambrone L, Pannuti CM, Guglielmetti MR, Chambrone LA. Evidence grade associating periodontitis with preterm birth and/or low birth weight: II. A systematic review of randomized trials evaluating the effects of periodontal treatment. *J Clin Periodontol*. 2011;38:902–14.
- Cho MK, Bero LA. Instruments for assessing the quality of drug studies published in the medical literature. *JAMA*. 1994;272:101–4.
- Contreras A, Herrera JA, Soto JE, Arce RM, Jaramillo A, Botero JE. Periodontitis is associated with preeclampsia in pregnant women. *J Periodontol*. 2006;77:182–8.
- Cochrane handbook for systematic reviews of interventions 4.2.5, Higgins JPT, Green S, editors. The Cochrane Library, Issue 3. Chichester, UK: John Wiley & Sons, Ltd; 2005. [Updated May 2005].
- Dechartres A, Charles P, Hopewell S, Ravaud P, Altman DG. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol*. 2011;64:136–44.
- Delaney A, Bagshaw SM, Ferland A, Laupland K, Manns B, Doig C. The quality of reports of critical care meta-analyses in the Cochrane Database of Systematic Reviews: an independent appraisal. *Crit Care Med*. 2007;35:589–94.
- Den Hartog L, Slater JJ, Vissink A, Meijer HJ, Raghoobar GM. Treatment outcome of immediate, early and conventional single-tooth implants in the aesthetic zone: a systematic review to survival, bone level, soft-tissue, aesthetics and patient satisfaction. *J Clin Periodontol*. 2008;35:1073–86.
- Detsky AS, Naylor CD, O'Rourke K, et al. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol*. 1992;45:255–65.
- Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309(6964):1286–91.
- Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health*. 1998;52:377–84.
- Sullivan BM. EBP@NUHS Ch 2: asking a good clinical question using the structured PICO format. 2008;2.3. <http://www.nuhs.edu/media/25458/studyguide-constructingsearchableclinicalquestion.pdf>
- Egger M, Davey Smith G. Meta-analysis: potentials and promise. *BMJ*. 1997;315:1371–4.
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–34.
- Egger M, Smith GD, Sterne JA. Uses and abuses in meta-analysis. *Clin Med*. 2001;1:478–84.
- Faggion CM Jr, Listl S, Giannakopoulos NN. The methodological quality of systematic reviews of animal studies in dentistry. *Vet J*. 2012;192:140–147.
- Faggion Jr CM, Schmitter M. Using the best available evidence to support clinical decisions in implant dentistry. *Int J Oral Maxillofac Implants*. 2010;25:960–9.
- Farquhar C, Vail A. Pitfalls in systematic reviews. *Curr Opin Obstet Gynecol*. 2006;18:433–9.
- Furlan AD, Brosseau L, Imamura M, Irvin E. Massage for low-back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine (Phila Pa 1976)*. 2002;27:1896–910.
- Garritty C, Tsertsvadze A, Tricco AC, Sampson M, Moher D. Updating systematic reviews: an international survey. *PLoS One*. 2010;5:e9914.
- Gilbody SM, Song F, Eastwood AJ, Sutton A. The causes, consequences and detection of publication bias in psychiatry. *Acta Psychiatr Scand*. 2000;102:241–9.
- Glenny AM, Esposito M, Coulthard P, Worthington HV. The assessment of systematic reviews in dentistry. *Eur J Oral Sci*. 2003;111:85–92.
- Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, et al. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med*. 2005;61:417–30.
- Greenhalgh T, Wong G, Westhorp G, Pawson R. Protocol – realist and meta-narrative evidence synthesis: evolving standards (RAMESES). *BMC Med Res Methodol*. 2011;11:115.
- Haroony M, Phillips R. “There is nothing like looking, if you want to find something” – asking questions and searching for answers – the evidence based approach. *Arch Dis Child Educ Pract Ed*. 2010;95:34–9.
- Haynes RB. Of studies, syntheses, synopses, summaries, and systems: the “5 S” evolution of information services for evidence-based healthcare decisions. *Evid Based Med*. 2006;11:162–4.
- Herrera D, Sanz M, Jepsen S, Needleman I, Roldán S. A systematic review on the effect of systemic antimicrobials as an adjunct to scaling and root planing in periodontitis patients. *J Clin Periodontol*. 2002;29 Suppl 3:136–59; discussion 160–2.
- Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;237:557–60.
- Higgins JPT, Altman DG. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: www.cochrane-handbook.org.
- Higgins JPT, Deeks JJ. Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: www.cochrane-handbook.org.
- Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: www.cochrane-handbook.org.
- Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify

- reports of randomized trials. *Cochrane Database Syst Rev.* 2007;(2):MR000001.
- Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database Syst Rev.* 2011;(8):MR000026.
- Hujoel PP. Endpoints in periodontal trials: the need for an evidence-based research approach. *Periodontol* 2000. 2004;36:196–204.
- Hujoel PP, DeRouen TA. A survey of endpoint characteristics in periodontal clinical trials published 1988–1992, and implications for future studies. *J Clin Periodontol.* 1995;22:397–407.
- Imperiale TF, McCullough AJ. Do corticosteroids reduce mortality from alcoholic hepatitis? A meta-analysis of the randomized trials. *Ann Intern Med.* 1990;113:299–307.
- Ioannidis JP, Lau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. *Proc Natl Acad Sci USA.* 2001;98:831–6.
- Ioannidis JP, Contopoulos-Ioannidis DG, Lau J. Recursive cumulative meta-analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. *J Clin Epidemiol.* 1999;52:281–91.
- Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol.* 1996;49:235–43.
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17:1–12.
- Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, Moher D. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *JAMA.* 1998;28:278–80.
- Jeffcoat MK, Hauth JC, Geurs NC, Reddy MS, Cliver SP, Hodgkins PM, Goldenberg RL. Periodontal disease and preterm birth: results of a pilot intervention study. *J Periodontol.* 2003;74:1214–8.
- Jeffcoat M, Parry S, Sammel M, Clothier B, Catlin A, Macones G. Periodontal infection and preterm birth: successful periodontal therapy reduces the risk of preterm birth. *BJOG.* 2011;118:250–6.
- Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA.* 1999;282:1054–60.
- Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ.* 2001;323(7303):42–6.
- Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol.* 2004;4:22.
- Khader YS, Jibreal M, Al-Omiri M, Amarín Z. Lack of association between periodontal parameters and preeclampsia. *J Periodontol.* 2006;77:1681–7.
- Koch GG, Paquette DW. Design principles and statistical considerations in periodontal clinical trials. *Ann Periodontol.* 1997;2:42–63.
- Koretz RL. Methods of meta-analysis: an analysis. *Curr Opin Clin Nutr Metab Care.* 2002;5:467–74.
- Kunnen A, Blaauw J, van Doormaal JJ, van Pampus MG, van der Schans CP, Aarnoudse JG, van Winkelhoff AJ, Abbas F. Women with a recent history of early-onset pre-eclampsia have a worse periodontal condition. *J Clin Periodontol.* 2007;34:202–7.
- Kunnen A, van Doormaal JJ, Abbas F, Aarnoudse JG, van Pampus MG, Faas MM. Periodontal disease and pre-eclampsia: a systematic review. *J Clin Periodontol.* 2010;37:1075–87.
- Lang S, Kleijnen J. Quality assessment tools for observational studies: lack of consensus. *Int J Evid Based Healthc.* 2010;8:247.
- Larue EM, Draus P, Klem ML. A description of a web-based educational tool for understanding the PICO framework in evidence-based practice with a citation ranking system. *Comput Inform Nurs.* 2009;27:44–9.
- Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med.* 1992;327:248–54.
- Lau J, Schmid CH, Chalmers TC. Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. *J Clin Epidemiol.* 1995;48:45–57.
- Law M. Evidence-based rehabilitation: a guide to practice. Thorofare: Slack Inc; 2002.
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* 2009;6:e1000100.
- List T, Axelsson S. Management of TMD: evidence from systematic reviews and meta-analyses. *J Oral Rehabil.* 2010;37:430–51.
- Lohsoonthorn V, Kungsadalpipob K, Chanchareonsook P, Limpongsanurak S, Vanichjakvong O, Sutdhibhisal S, Sookprom C, Wongkittikraiwan N, Kamolpornwiji W, Jantarasaengaram S, Manotaya S, Siwawej V, Barlow WE, Fitzpatrick AL, Williams MA. Maternal periodontal disease and risk of preeclampsia: a case-control study. *Am J Hypertens.* 2009;22:457–63.
- Loos BG, Tjoa S. Host-derived diagnostic markers for periodontitis: do they exist in gingival crevice fluid? *Periodontol* 2000. 2005;39:53–72.
- López NJ, Smith PC, Gutierrez J. Periodontal therapy may reduce the risk of preterm low birth weight in women with periodontal disease: a randomized controlled trial. *J Periodontol.* 2002;73:911–24.
- Lu CY. Observational studies: a review of study designs, challenges and strategies to reduce confounding. *Int J Clin Pract.* 2009;63:691–7.
- Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med.* 2002;21:2313–24.
- Lundh A, Knijnenburg SL, Jørgensen AW, van Dalen EC, Kremer LC. Quality of systematic reviews in pediatric oncology – a systematic review. *Cancer Treat Rev.* 2009;35:645–52.
- Lutje V, MacLehose H, Garner P. Editorial strategy for revising Cochrane reviews: does it help? In: 13th Cochrane colloquium in Melbourne; 2005. p. 153.
- Macones GA, Parry S, Nelson DB, Strauss JF, Ludmir J, Cohen AW, Stamilio DM, Appleby D, Clothier B,

- Sammel MD, Jeffcoat M. Treatment of localized periodontal disease in pregnancy does not reduce the occurrence of preterm birth: results from the Periodontal Infections and Prematurity Study (PIPS). *Am J Obstet Gynecol*. 2010;202:147.e1–8.
- Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther*. 2003;83:713–21.
- Manchikanti L, Benyamin RM, Helm S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 3: systematic reviews and meta-analyses of randomized trials. *Pain Physician*. 2009;12:35–72.
- Meurman JH, Furuholm J, Kaaja R, Rintamaki H, Tikkanen U. Oral health in women with pregnancy and delivery complications. *Clin Oral Investig*. 2006;10:96–101.
- Michalowicz BS, Hodges JS, DiAngelis AJ, Lupo VR, Novak MJ, Ferguson JE, Buchanan W, Bofill JJ, Papanou PN, Mitchell DA, Matseoane S, Tschida PA, OPT Study. Treatment of periodontal disease and the risk of preterm birth. *N Engl J Med*. 2006;355:1885–94.
- Miller SA, Miller G. Practice for emergency treatment of dental trauma: EB case report. *J Evid Base Dent Pract*. 2010;10:135–46.
- Moher D, Tsartsvadze A. Systematic reviews: when is an update an update? *Lancet*. 2006;367:881–3.
- Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials*. 1995;16:62–73.
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet*. 1999;354:1896–900.
- Moher D, Tsartsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, Barrowman N. A systematic review identified few methods and strategies describing when and how to update systematic reviews. *J Clin Epidemiol*. 2007;60(11):1095–104.
- Moher D, Tsartsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, Barrowman N. When and how to update systematic reviews. *Cochrane Database Syst Rev*. 2008;(1):MR000023.
- Moles DR, Needleman IG, Niederman R, Lau J. Introduction to cumulative meta-analysis in dentistry: lessons learned from undertaking a cumulative meta-analysis in periodontology. *J Dent Res*. 2005;84:345–9.
- Montenegro R, Needleman I, Moles D, Tonetti M. Quality of RCTs in periodontology – a systematic review. *J Dent Res*. 2002;81:866–70.
- Mullen B, Muerlreile P, Bryant B. Cumulative meta-analysis: a consideration of indicators of sufficiency and stability. *Pers Soc Psychol Bull*. 2001;27:1450–62.
- Nabet C, Lelong N, Colombier ML, Sixou M, Musset AM, Goffinet F, Kaminski M. Maternal periodontitis and the causes of preterm birth: the case-control Epipap study. *J Clin Periodontol*. 2010;37:37–45.
- National Institute of Dental and Craniofacial Research (NIDCR). U.S. Department of Health and Human Services, National call to action to promote oral health in America: a report of the surgeon general. 2003. Available at: www.surgeongeneral.gov/topics/oral-health/nationalcalltoaction.htm. Accessed July 2011.
- Needleman IG. A guide to systematic reviews. *J Clin Periodontol*. 2002;29 Suppl 3:6–9.
- Needleman I, Moles DR, Worthington H. Evidence-based periodontology, systematic reviews and research quality. *Periodontol* 2000. 2005;37:12–28.
- Newnham JP, Newnham IA, Ball CM, Wright M, Pennell CE, Swain J, Doherty DA. Treatment of periodontal disease during pregnancy: a randomized controlled trial. *Obstet Gynecol*. 2009;114:1239–48.
- Nguyen QV, Bezemer PD, Habets L, Prahl-Andersen B. A systematic review of the relationship between overjet size and traumatic dental injuries. *Eur J Orthod*. 1999;21:503–15.
- Oettinger-Barak O, Barak S, Ohel G, Oettinger M, Kreutzer H, Peled M, Machtei EE. Severe pregnancy complication (preeclampsia) is associated with greater periodontal destruction. *J Periodontol*. 2005;76:134–7.
- Offenbacher S, Lin D, Strauss R, McKaig R, Irving J, Barros SP, Moss K, Barrow DA, Hefti A, Beck JD. Effects of periodontal therapy during pregnancy on periodontal status, biologic parameters, and pregnancy outcomes: a pilot study. *J Periodontol*. 2006;77:2011–24.
- Offenbacher S, Beck JD, Jared HL, Mauriello SM, Mendoza LC, Couper DJ, Stewart DD, Murtha AP, Cochran DL, Dudley DJ, Reddy MS, Geurs NC, Hauth JC, Maternal Oral Therapy to Reduce Obstetric Risk (MOTOR) Investigators. Effects of periodontal therapy on rate of preterm delivery: a randomized controlled trial. *Obstet Gynecol*. 2009;114:551–9.
- Oliveira AM, de Oliveira PA, Cota LO, Magalhães CS, Moreira AN, Costa FO. Periodontal therapy and risk for adverse pregnancy outcomes. *Clin Oral Investig*. 2010. doi:10.1007/s00784-010-0424-8.
- Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess the quality of randomized controlled trials: a systematic review. *Phys Ther*. 2008;88(2):156–75.
- Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol*. 1991;44:1271e8.
- Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, Streiner DL. Agreement among reviewers of review articles. *J Clin Epidemiol*. 1991;44:91–8.
- Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review – a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy*. 2005;10:21–34.
- Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, Tharyan P, Colford JM Jr. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. *Natl Med J India*. 2004;17:86–95.
- Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Control Clin Trials*. 1997;18:580–93.
- Paradis C. Bias in surgical research. *Ann Surg*. 2008;248:180–8.

- Radnai M, Pál A, Novák T, Urbán E, Eller J, Gorzó I. Benefits of periodontal therapy when preterm birth threatens. *J Dent Res*. 2009;88:280–4.
- Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics*. 1989;84:815–27.
- Richardson W, Wilson M, Nishikawa J, Hayward R. The well-built clinical question: a key to evidence-based decisions. *ACP J Club*. 1995;123:A12–3.
- Ried K. Interpreting and understanding meta-analysis graphs – a practical guide. *Aust Fam Physician*. 2006;35:635–8.
- Rosenthal R. The “file drawer problem” and tolerance for null results. *Psychol Bull*. 1979;86:638–41.
- Rochon PA, Gurwitz JH, Sykora K et al. Reader’s guide to critical appraisal of cohort studies: 1. Role and design. *BMJ* 2005;330:895–7.
- Sackett DL. Evidence-based medicine: how to practice and teach EBM. 2nd ed. Edinburgh/New York: Churchill Livingstone; 2000, xiv, 261.
- Sadatmansouri S, Sedighpoor N, Aghaloo M. Effects of periodontal treatment phase I on birth term and birth weight. *J Indian Soc Pedod Prev Dent*. 2006;24:23–6.
- Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res*. 2008;17:279–301.
- Salmos J, Gerbi ME, Braz R, Andrade ES, Vasconcelos BC, Bessa-Nogueira RV. Methodological quality of systematic reviews analyzing the use of laser therapy in restorative dentistry. *Lasers Med Sci*. 2010;25:127–36.
- Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*. 2007;36:666–76.
- Santos CMC, Pimenta CAM, Nobre MRC. The PICO strategy for the research question construction and evidence search. *Rev Latino-am Enfermagem*. 2007;15:508–11.
- Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak*. 2007;7:16.
- Scherer S, Smith MB. Teaching evidence-based practice in academic and clinical settings. *Cardiopulm Phys Ther*. 2002;13:23.
- Shamliyan T, Kane RL, Dickinson S. A systematic review of tools used to assess the quality of observational studies that examine incidence or prevalence and risk factors for diseases. *J Clin Epidemiol*. 2010;63:1061–70.
- Shea B, Moher D, Graham I, Pham B, Tugwell P. A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. *Eval Health Prof*. 2002;25:116–29.
- Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7:10.
- Shetty M, Shetty PK, Ramesh A, Thomas B, Prabhu S, Rao A. Periodontal disease in pregnancy is a risk factor for preeclampsia. *Acta Obstet Gynecol Scand*. 2010;89:718–21.
- Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med*. 2007a;147:224–33.
- Shojania KG, Sampson M, Ansari MT, Ji J, Garrity C, Rader T, Moher D. Updating systematic reviews. Rockville: Agency for Healthcare Research and Quality (US); 2007b. Report No.: 07–0087.
- Sindhu F, Carpenter L, Seers K. Development of a tool to rate the quality assessment of randomized controlled trials using a Delphi technique. *J Adv Nurs*. 1997;25:1262–8.
- Siqueira FM, Cota LO, Costa JE, Haddad JP, Lana AM, Costa FO. Maternal periodontitis as a potential risk variable for preeclampsia: a case–control study. *J Periodontol*. 2008;79:207–15.
- Smith LA, Oldman AD, McQuay HJ, Moore RA. Teasing apart quality and validity in systematic reviews: an example from acupuncture trials in chronic neck and back pain. *Pain*. 2000;86:119–32.
- Souza JP, Pileggi C, Cecatti JG. Assessment of funnel plot asymmetry and publication bias in reproductive health meta-analyses: an analytic survey. *Reprod Health*. 2007;4:3.
- Srinivas SK, Sammel MD, Stamilio DM, Clothier B, Jeffcoat MK, Parry S, Macones GA, Elovitz MA, Metlay J. Periodontal disease and adverse pregnancy outcomes: is there an association? *Am J Obstet Gynecol*. 2009;200:497.e1–8.
- Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283:2008–12.
- Suebnuksarn S, Ngamboonsirisingh S, Rattanabanlang A. A systematic evaluation of the quality of meta-analyses in endodontics. *J Endod*. 2010;36:602–8.
- Tarannum F, Faizuddin M. Effect of periodontal therapy on pregnancy outcome in women affected by periodontitis. *J Periodontol*. 2007;78:2095–103.
- The Cochrane Collaboration. Maintaining your review. The Cochrane Collaboration open learning material; 2002. Module 19.
- van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine*. 2003;28:1290–9.
- Verhagen AP, de Vet HC, de Bie RA, et al. Balneotherapy and quality assessment: interobserver reliability of the Maastricht criteria list and the need for blinded quality assessment. *J Clin Epidemiol*. 1998a;51:335–41.
- Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol*. 1998b;51:1235–41.
- Wells GA, Shea B, O’Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised

- studies in meta-analyses. University of Ottawa. 2001. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm. Accessed 17 Oct 2011.
- West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence, evidence report, Technology Assessment No. 47. AHRQ Publication No. 02-E016. Rockville: Agency for Health Care Research and Quality; 2002. www.thecre.com/pdf/ahrq-system-strength.pdf.
- Wilczynski NL, Haynes RB; Hedges Team. Robustness of empirical search strategies for clinical content in MEDLINE. *Proc AMIA Symp.* 2002;904–8. <http://www.ncbi.nlm.nih.gov/pubmed/12463956>
- Westbrook JI, Ampt A, Williamson M, Nguyen K, Kearney L. Methods for measuring the impact of health information technologies on clinicians' patterns of work and communication. *Stud Health Technol Inform.* 2007;129(Pt 2):1083–7.
- Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Med Res Methodol.* 2010;10:54.
- Yates SL, Morley S, Eccleston C, de C Williams A. A scale for rating the quality of psychological trials for pain. *Pain.* 2005;117:314–25.

Over the past 30 years, the periodontal literature has become more evidence based. During the transition, it is not surprising that investigators have come to rely increasingly on statistically based probability statements to validate their investigative claims. However, it seems fair to claim that most experienced investigators would now agree that even appropriate statistical analyses are not panaceas—if for no other reason than that such analyses are only as valid as the assumptions which underlie the statistical tests that are used. The point of this chapter so far is to help readers understand that the newly popular debate about clinical significance versus statistical significance is not alone among the concerns they should have regarding the clinical meaning of statistically based scientific literature, as a claim of the clinical insignificance of a statistically significant result depends on a critical assessment of the sometimes unique context in which a clinical decision must endure (Rethman and Nunn 1999).

9.1 The Matter of Statistical Significance

There are four steps in the development of an analytical investigation: establishing the research hypothesis (association between exposure and outcomes believed to exist), study design, data collection, and statistical assessments that include hypothesis testing. Statistical significance testing, also called **hypothesis testing**, is an exten-

sion of the scientific method. It provides a yes or no answer to the question of whether there are differences between variables, or outcome measures, in test and control groups. The term ‘**statistical significance**’, however, only denotes that the associations between tested variables did not occur by chance. The term ‘statistically significant at an α level of 0.05’ means that the null hypothesis (that there is no relationship regarding specific variables between test and control groups) was rejected and the chances of the association occurring by chance was small (probability is 5% or less, or the odds are 1 out of 20) (Greenstein 2003a).

When making decisions about whether to reject a null hypothesis, we can make two types of error (**Table 9.1**):

- (a) Declaring that there is a difference between the groups, when in fact there is no difference (rejecting the null hypothesis when it is in fact true; type I error)
- (b) Deciding there is no difference, when in fact a difference exists (type II error) (Christley 2008a)

Issues related to the inability of hypothesis testing to provide clarity regarding detecting clinically relevant periodontal changes follow. Failure to reject the null hypothesis does not mean that no true difference exists between the test and control groups. Chance or too few subjects being enrolled in the study can prevent the finding of a statistically significant difference when one exists. Researchers can avoid the problem of including too few subjects by performing a **power**

Table 9.1 Possible outcomes from testing experimental hypotheses (Christley2008a) (Reprinted with permission from John Wiley & Sons, Inc.)

Result of statistical test	Reality	
	Null hypothesis is true	Null hypothesis is false
Reject null hypothesis	Type I error probability = significance level	Correct conclusion probability = power
Do not reject null hypothesis	Correct conclusion	Type II error probability = $1 - \text{power}$

analysis to determine the appropriate sample size that needs to be evaluated (Greenstein 2003a).

There are **three types of power analysis** viz. a priori, post hoc, and compromise (Singh 2006):

- **A priori power analysis** is carried out during the design stage of the study for determining sample size for a given significance level, power, and effect size. The effect size may be chosen based on substantive knowledge, previous research or by conventions. An a priori power analysis ensures savings in time and resources while conducting a study which has a little chance of finding a significant effect and also restricts testing more subjects than are necessary to detect an effect.
- **Post hoc power analysis** is done after a study has been carried out for determining power of the test for given sample size, significance level, and the observed effect size. This helps in explaining the results of a study that did not find any significant effect.
- **Compromised power analysis** may be carried out in situations where pragmatic constraints restrict adherence to the recommendations derived from a priori power analysis. The maximum possible sample size and the effect size are fixed, and alpha and beta error probabilities are chosen in accordance with relative seriousness.

Obviously, a priori power analysis is more useful than a post hoc one, whereas compromised power analysis is more complex, is rarely used, and constitutes controversial issue (Singh 2006).

Further, power analysis also has a role in ethical issues. If a study to test a new drug will have required power with a sample of 50 patients, then it would be inappropriate to use a sample of 100 patients since the other 50 are being unnecessarily put at risk. Similarly, if a study requires 100

patients to yield desirable power, it would be inappropriate to use only 50 because these 50 patients have been put at risk for no reason (Singh 2006).

The power of a study is the degree to which we are certain that if we conclude a difference does not exist, it in fact does not exist. This is determined by $1 - \beta$ (power = $1 - \beta$). Since β is generally set at 0.2, the accepted level of power is $1 - 0.2 = 0.8$ or 80%. The way to ensure a power of 80% is to do a sample size calculation, which answers the question, ‘If I want to be 95% certain that any difference I see is not due to chance and 80% certain that if I conclude there is no difference I am correct, how many people do I need in this study?’ (Godwin 2001).

If an article concludes no difference was found, the authors should tell you the level of certainty (power) with which they can make that conclusion. Of course, if a statistical difference is seen ($P < 0.05$ or the 95% CI does not include 1), then by definition there was sufficient power. If the power is really high (i.e. if there is a huge sample size compared with the number actually needed so that the power is, say, 99%), statistical differences can be seen even when very small real clinical differences exist. For instance, an RR of 1.2 with a 95% CI of 1.1–1.3 might be highly statistically significant; the CI is very narrow due to the large sample size. This means the difference is likely to be real and not due to chance, but is the difference clinically significant? Sometimes it is, depending on the seriousness of the issue. If the RR indicates a child is 1.2 times more likely to die within the next 3 months if exposed to X, it is highly clinically significant. If it indicates that people are 1.2 times more likely to get a runny nose if they go out in cold weather without a hat, it is less important (Godwin 2001).

Determination of power of the test is required when the test results in non-significance. It is of great value in those situations also in which statistical significance may bear little relation to clinical significance and a conventional analysis using *P* values is liable to be misleading (Singh 2006).

Therefore, there are four key parameters that we should consider when planning a study (Christley 2008b):

- The minimum difference we wish to detect (this may be based on clinical knowledge; in our example, we may only be interested in detecting a difference in blood pressure of 10 mmHg or greater)
- The type I error rate (this is almost always set at 0.05)
- The type II error rate (usually set at 0.2; i.e. power of 80%)
- The sample size

These parameters are interlinked—a certain sample size is needed to detect a particular minimum difference, with type I and II error rates of 0.05 and 0.2, respectively. If we wish to increase the power above 80% without affecting the minimum detectable difference or type I error rate, we need to increase the sample size. If we have fewer animals in our study than the required sample size (and we keep the type I error rate at 0.05), we are at increased risk of missing an important difference, either because we can only detect a difference greater than the minimum we thought clinically important, or because the type II error rate has increased (Christley 2008b).

The term ‘**effect size**’ refers to the magnitude of the effect under the alternate hypothesis (Singh 2006). The bigger it is, the easier it will be to find. The nature of effect size varies from one statistical procedure to the other, though its function in power analysis remains same in all the procedure. For example, in paired *t* test, the effect size is the mean of difference divided by standard deviation of difference, while in unpaired *t* test, effect size is the absolute difference between two group means divided by pooled standard deviation. The effect size for correlation coefficient *r* is simply the *r* itself. Cohen (1988) has defined small, medium, and large effect sizes for many types of tests that may form useful conventions. For unpaired *t* test, the conven-

tional small, medium, and large effect sizes are 0.20, 0.50, and 0.80, respectively, whereas for paired *t* test, these are 0.10, 0.25, and 0.40, respectively. Conventional small, medium, and large effect sizes in case of correlation coefficient are 0.10, 0.30, and 0.50, respectively (Singh 2006).

9.2 Definition of Clinical Significance

It sometimes seems that demonstration of ‘statistical significance’ has become paramount in the minds of many researchers; indeed this is sometimes deemed by authors, reviewers, and editors alike as a prerequisite to publication (Table 9.2). Clearly, statistical and clinical significance are related since, with few exceptions, a non-statistically significant finding is unlikely to have clinical significance other than to rule out clinical relevance. Exceptions could be in proving equivalence of an outcome to an already established gold standard, for example, a chemical plaque control agent with the same plaque inhibitory properties as chlorhexidine. The clinical significance would be further enhanced if the new agent also had fewer side effects than the gold standard and/or was less expensive. Statistical significance, conversely, does not necessarily equate with clinical significance. The level of probability of a finding is not a guide to clinical significance but merely relates to the chance that the finding is due to a type I or alpha error (Addy and Newcombe 2005; Hollon and Flick 1988; Kingman 1992; LeFort 1993; Lindgren et al. 1993).

As a corollary, some readers may perceive that demonstration of a statistically significant difference, for example, in the effect of two drugs, is the same as demonstrating the existence of an important difference, and that failure to demonstrate a significant difference indicates that no difference exists. Such conclusions are misleading and often incorrect (Christley 2008a).

When considering the results of the study, we should ask:

- (a) Is the observed difference between the groups real or could it be a chance finding?
- (b) If the difference is real, is it large enough to be important? (Christley 2008a)

Table 9.2 Various definitions of clinical significance found in the literature

Author and year	Definition of clinical significance
Hujoel et al. (2000)	“statistically significant difference in a clinically important outcome identified in a definitive or phase III clinical trial”
Killooy (2002)	“clinical significance is a subjective evaluation of significance by the clinician and that before a finding can be clinically significant, it must have achieved statistical significance”
Greenstein (2003a)	“clinical significance denotes a change that may alter how a clinician will treat a patient, and this value judgment varies depending on the situation. Clinicians, researchers, industry representatives, third-party payers and so forth may interpret clinical relevance differently, as they may place emphasis on different outcomes (for example, size of effect, cost, time needed for therapy, ease of implementation, duration of results and consumer acceptability). To arrive at a conclusion that a result is clinically significant, the finding must be clinically meaningful and statistically significant”
Addy and Newcombe (2005)	“To appraise both the distinction and relationship between the two concepts, clinical and statistical significance, it would be valuable to have a clear definition, or at least a description, of both. This is relatively straightforward for statistical significance but not for clinical significance, for which a description might be possible but for which there is no agreed definition. Statistical significance is a mathematical concept based on hypothesis testing and in published literature provides evidence that conclusions drawn by the author are not well explained by chance. Clinical significance in periodontology is a much more nebulous concept for which there are no agreed rules”

Statistical tests can only help directly with the first of these questions; interpretation of the clinical importance of the results (question b) requires clinical/biological judgement and cannot be assessed solely through statistical testing (Christley 2008a).

9.3 Methods and Criteria to Evaluate Clinical Relevance

With the approval of interventions or treatments for periodontal disease based on results of large clinical trials, it is often difficult for the clinician to determine whether the statistical results obtained from these studies are clinically significant and whether these interventions or treatments will thus be effective in their individual patients. A series of measures of clinical significance derived from the literature and a 2-step approach was suggested by Van Dyke (2005) for determining clinical significance. First, the reliability of the data as a function of the reliability of the measurement tool(s) is calculated, and then clinical significance is put into the context of effect

size. By using these estimators, clinicians can determine, for example, whether administration of an experimental local medication plus scaling and root planing provides a clinically significant outcome, which can assist them in deciding which interventions are most effective for their patients (Van Dyke 2005).

Clinical significance, however, poses a dilemma. Clinical significance is a ‘subjective’ evaluation of significance by the clinician. Before a finding can be clinically significant, it must have achieved statistical significance or else it could be a chance happening. The clinician should base his decision regarding clinical significance on some reasonable criteria. The magnitude of change could be clinically significant if it simplifies (complicates) the nature of future treatment, if it reduces the need for more aggressive therapy, if it makes maintenance easier for the patient and clinician, or if it makes restorative dentistry easier for the patient and clinician. The magnitude of change could be clinically significant if it improves (worsens) the immediate or long-term prognosis (e.g. ranging from hopeless to retainable, poor to fair, or fair to good) and if it improves the classification of furcation (from III to II, or II to I). Other factors

Table 9.3 Periodontal index of clinical significance: Potential outcomes that could be considered^a (Greenstein and Lamster 2000) (Reproduced with permission from the American Academy of Periodontology)

Direct measures
Clinical assessments
• Mean changes in probing depth in millimetres and clinical attachment level in millimetres for the patient • Percentage of sites demonstrating a 2-mm reduction in probing depth or clinical attachment gain, or CAG. Other thresholds also could be used^{b,c} • Percentage of sites (or patients) with initially deep probing depth (>5 mm) that were reduced to 5 mm or less^{b,c} • Percentage of sites demonstrating 1 mm, 2 mm and so forth reduction in probing depth or CAG with regard to various initial probing depths^{b,c} • Percentage improvement of any parameter considering the severity of the defect being treated • Reduction of the incidence of disease progression (threshold for disease progression being defined as a 2-mm or more loss of clinical attachment) • Number of sites (or patients) that need to be treated to achieve or prevent an event • Percentage of inflamed sites (for example, bleeding on probing) in which clinical inflammation was eliminated^{b,c} • Persistence of improvement observed after therapy over defined periods (measured in months or years)
Radiographic assessments
• Percentage of bone fill in osseous defects
Radiographic assessments
• Percentage of bone fill in osseous defects
Indirect measures
• Percentage of sites (or patients) requiring additional therapy after treatment^{b,c} • Cost-benefit ratio • Length of therapy • Risks associated with new therapies • Patient satisfaction

^aAll measures are compared to a standard of care (for example, scaling and root planing or conventional flap surgery)

^bPercentages should be accompanied by actual numbers (for example, number of sites)

^cThere are statistical techniques that take into account changes at multiple sites in one patient so they are not treated as independent events

that impact on clinical significance include time, cost, and morbidity (Killooy 2002).

Clinical attachment levels, bone height, and bone density are primary outcome variables commonly used to monitor periodontal patients. They are considered primary outcome variables since they are part of the tooth's attachment apparatus. Secondary outcome variables include probing depths, mobility patterns, bleeding on probing, and biochemical and microbiological assessments (Greenstein 2003a). Clinical trials are conducted to answer clinical questions, and clinical parameters are used to monitor outcomes; therefore, the results should refer to the importance of the clinical data. Identifying specific criteria related to relevant clinical periodontal parameters could enhance interpretation of results and selection of the most suitable treatments for various clinical problems. If a consensus could be reached

about the definition of clinical significance and thresholds of specific clinical parameters that reflect clinically meaningful findings, then some of the criteria listed in the table could be incorporated into a list of determinants to assess clinical significance (Table 9.3) (Greenstein 2003a). The advantages and disadvantages of methods and criteria that could help define clinical significance were reviewed by Greenstein (2003a).

Subjective true endpoints such as 'do you notice bleeding after brushing?' are simpler and less costly to obtain than the objective surrogate endpoints such as gingival index scores. Yet, subjective true endpoints are almost never used in clinical trials because of the notion that objective surrogates are superior endpoints. This is unfortunate. The goal of therapy is to provide tangible patient benefits, not tangible clinician benefits. If the goal of free connective

tissue grafts is to improve aesthetics, then the self-reported aesthetics improvement, and not recession measurements, should be the primary outcome measurement in a clinical trial. Just as the primary endpoint in a trial on knee surgery was the subjective self-report of knee function (48) (not objective knee function tests, which are surrogates), so should the primary endpoints in periodontal trials be studied from the patient perspective, not the clinician perspective (Hujoel 2004).

The clinical relevance of surrogate endpoint changes can be imputed when, based on epidemiologic observations, surrogate endpoint changes can be translated into true endpoint changes. For instance, if blood pressure medication reduces the blood pressure by 10 mmHg, and if epidemiologic data suggest that such a blood pressure decrease is associated with a 15% drop in cardiovascular disease mortality risk, then that medication can be imputed to reduce cardiovascular disease mortality by 15%. Because of the uncertainty regarding the relationships between surrogate endpoints and true endpoints, such 'translations' are extremely dangerous, but may give some handle on clinical significance when epidemiologic data are available (Hujoel 2004).

In periodontal clinical trials, clinical attachment level measurements are commonly used as surrogates for tooth loss. However, it was provided evidence that moderate attachment losses were informative on tooth mortality. Hujoel et al. (1997) revealed that both 1 and 2 year changes in probing attachment level were informative about tooth mortality risk. A 1-mm loss measured over a 1-year period was associated with a 56% increased tooth mortality risk (relative risk = 1.56; 95% confidence interval, 1.08–2.26; $P=0.017$); a 1-mm loss measured over a 2-year period was associated with a 102% increased risk for tooth mortality (relative risk = 2.02; 95% confidence interval, 1.26–3.25; $P=0.004$). Both the lifetime cumulative attachment loss and attachment loss since young adulthood of ≥ 2 or ≥ 3 mm were informative on tooth mortality. Tooth mortality risk increased as the attachment loss increased; loss ≥ 3 mm at the buccal or mesial site increased

tooth mortality risk, by 91% (relative risk, 1.91; 95% confidence interval, 1.01–3.60) and 270% (RR, 3.70; 95% CI, 1.83–7.49), respectively (Hujoel et al. 1999).

It is the responsibility of each of us to evaluate the magnitude of any treatment effect. It is likely that the type of procedure will determine the magnitude and time course of change. Other variables, such as experimental design, severity of the disease at baseline, and confounders such as smoking, may change the magnitude of effect. Each clinician should decide if the magnitude of change will benefit each individual patient in light of these factors—the risk/benefit ratio and cost/benefit ratio. Every effort should be made to achieve this assessment with minimal bias (Jeffcoat 2002).

9.4 The 'Number Needed to Treat' (NNT)

The inverse of the difference between rates, called the 'number needed to treat' (NNT), was suggested 20 years ago as a good way to present the results of comparisons of success or failure under different therapies. Such comparisons usually arise in randomized controlled trials and meta-analysis (Hutton 2009).

9.4.1 Definition of NNT

The NNT represents the number of persons (it can also be done for the number of sites) that must be treated in the test group during a given time period to achieve one additional treatment result or to prevent one adverse incident compared to the control group (Greenstein 2006; Greenstein and Nunn 2004).

For the GTR meta-analysis, and using the benefit of proportion of sites achieving at least 2-mm gain in attachment, the NNT is eight. In other words, for every eight patients treated with GTR, you can expect one to have at least 2 mm more gain in clinical attachment than if you had used an access flap (Needleman et al. 2005).

Table 9.4 Relation between risk (of efficacy or harm) with treatment and control, relative risk, relative and absolute risk reduction, and number needed to treat. An alternative way to calculate the relative risk reduction is $1 - RR$ (Tramèr and Walder 2005) (Reprinted with permission from Springer Science+Business Media)

Risk with treatment EER	Risk with control CER	Relative risk	Relative risk reduction	Absolute risk reduction	Number needed to treat
0.2	0.5	0.4	0.6	0.3	3.3
0.02	0.05	0.4	0.6	0.03	33.0
0.002	0.005	0.4	0.6	0.003	333.0

EER experimental event rate, CER control event rate, RR relative risk, RRR relative risk reduction, ARR absolute risk reduction

9.4.2 Calculating the NNT

In randomized studies, patients are randomly allocated to an experimental intervention (that is thought to have some beneficial effects) and to a control intervention. In a placebo-controlled trial, the control intervention is a placebo; in an active-controlled trial, the control intervention is another active treatment. The incidence of an event occurring with the experimental intervention may be called the **experimental event rate (EER)**; accordingly, the incidence of the event occurring with the control intervention is the **control event rate (CER)** (Tramèr and Walder 2005).

The risk without therapy, also called **baseline risk** or **underlying risk**, may have a major impact on the effect of a treatment. For instance, when the underlying risk is extremely low, even an effective therapy has no scope to show efficacy, and when the underlying risk is extremely high, the efficacy of an intervention may be exaggerated. Therefore, to take into account the underlying risk in a study population, the **absolute risk reduction (ARR)** is often considered an additional measure of clinical effectiveness. The absolute risk reduction is the numerical difference between the control and experimental event rates ($ARR = CER - EER$). This number provides an idea about the clinical relevance of the effect of a treatment. However, absolute risk reduction is problematic as a measure of efficacy or harm, as it is a dimensionless, abstract number that may be difficult to incorporate into clinical practice. The number needed to treat (NNT), which is simply the reciprocal of the absolute risk reduction ($NNT = 1/ARR = 1/[CER - EER]$), has the advantages of the absolute risk reduction and

additionally provides a way of expressing the effect of a treatment in clinical terms. The number needed to treat is the number of patients who must be treated with an experimental intervention to achieve a particular result (beneficial or harmful) in one of them which would not have been the case had they all received the control intervention. **Table 9.4** summarizes the relations between experimental and control event rates, relative risk, relative risk reduction, absolute risk reduction, and number needed to treat. In analogy to the quantification of benefit (number needed to treat) and harm (**number needed to harm**), the **number needed to screen** has been proposed as a further use of the number needed to treat concept; this is the number of patients who must be screened for a disease to prevent one adverse outcome (Tramèr and Walder 2005; Hutton 2009).

If the results are presented as the percentage of times (mean per patient) an event occurred at sites, to determine the NNT, calculate the percentage difference in the event rates between the test and control groups and divide 100 by this number. This will indicate the average number of sites that need to be treated in the test group to attain an outcome at one additional location compared to the control group (Greenstein 2006).

The number needed to treat is a meaningful way of expressing the benefit of an active treatment over a control. It can be used either for summarizing the results of a therapeutic trial or for medical decision making about an individual patient, but its use at the bedside has been impeded by the need for time-consuming calculations. A nomogram has therefore been devised that will greatly simplify the calculations (**Fig. 9.1**). The

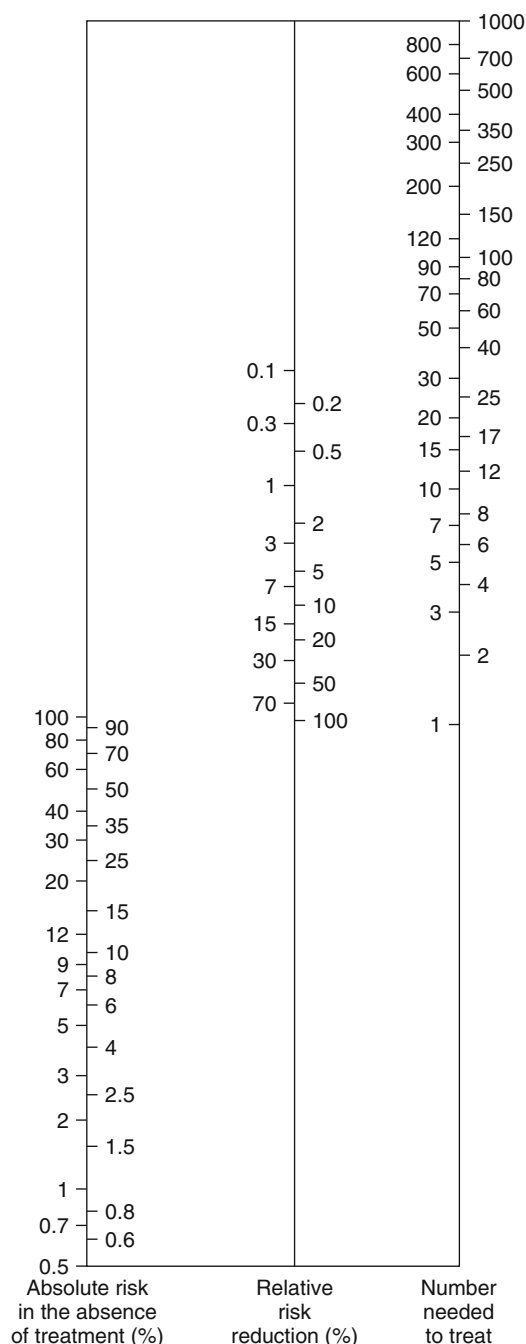


Fig. 9.1 Nomogram for calculating the number needed to be treated (Chatellier et al. 1996. Reprinted with permission from BMJ Publishing Group Ltd.)

nomogram allows the number needed to treat to be obtained directly without any calculation: a straight line should be drawn from the point

corresponding to the proportion of events in the control group on the left-hand scale to the point corresponding to the relative risk reduction measured in the trial on the central scale. The point of intercept of this line with the right-hand scale gives the number needed to treat. By taking the upper and lower limits of the confidence interval of the relative risk reduction, we can then obtain the upper and lower limits of the number needed to treat. This allows us to assess the precision of the result and the magnitude of effectiveness on the most optimistic and the most pessimistic hypotheses (Chatellier et al. 1996).

For detailed guidance regarding the use and calculation of the NNT, the reader is recommended to the electronic journal 'Bandolier': <http://www.jr2.ox.ac.uk/bandolier/booth/painpg/NNTstuff/numeric.htm>.

9.4.3 Requirements for Use of NNT

When calculating the NNT, the following items should be clearly denoted: treatment technique, specific therapeutic outcome to be assessed, duration of treatment that is necessary to achieve the outcome, and length of the observation period. The NNT can be used to compare procedures if the therapy is directed at the same condition and the same measurements are employed over a similar period of time (Greenstein and Nunn 2004).

9.4.4 Interpretation of NNT

An NNT of 1 indicates that a desirable result occurs in almost every patient or site that the treatment is administered. An NNT of 2 or 3 denotes that the treatment is very efficacious (Greenstein and Nunn 2004).

9.4.5 Benefits of Calculating NNT

1. Numerical and statistical data can be interpreted in terms that can be directly applied to patients, because it shows the effort required to achieve a particular therapeutic target.

Table 9.5 Advantages and limitations of employing NNT calculations (Greenstein and Nunn 2004) (Reproduced with permission from the American Academy of Periodontology)**Benefits of calculating the NNT**

- It shows the effort required to achieve a particular therapeutic target
- It can be applied to any beneficial outcome or adverse event
- Estimating the NNT for numbers needed to harm (e.g., side effects) can provide important information regarding the benefit/risk ratio of a therapy
- The NNT can be a guide to the overall net value of a prophylactic intervention with regard to an individual's risk (e.g., future periodontal disease progression)

Limitations of employing the NNT

- NNT is expressed as a single number, which is known as a point estimate. The actual result can be higher or lower than a point estimate. The 95% confidence interval can be supplied
- The NNT expresses frequency that an event will be achieved or prevented; it does not denote that this event is important
- When calculating the NNT, patients with dissimilar risk of periodontal disease progression should not be assessed together (e.g., different diseases or degrees of disease severity)
- The NNT tells you nothing about the health status of individuals who did not benefit from the treatment
- NNT is valid for comparing therapies. However, the studies need to have been monitored over a similar span of time and same types of measurements need to be employed
- NNT values assume that effective therapy is distributed equally. However, means from groups of patients are used to calculate NNT values. Therefore, NNT values may not reflect very successful therapy in some subjects

2. It can be applied to any beneficial outcome or adverse event if the appropriate data are available for evaluation.
3. Estimating the NNT for numbers needed to harm (e.g. side effects) can provide important information regarding the benefit/risk ratio of a therapy.
4. The NNT can be guide to the overall net value of a prophylactic intervention with regard to an individual's risk (e.g. future periodontal disease progression) (Greenstein and Nunn 2004).

9.4.6 Limitations of Employing NNT

Although NNTs are powerful instruments for interpreting clinical effects, they also have important limitations (McQuay and Moore 1997; Greenstein and Nunn 2004) (Table 9.5).

Needleman et al. (2002) systematically reviewed the evidence for efficacy of guided tissue regeneration (GTR) for infrabony defects. A *post hoc* analysis of GTR showed a significant benefit for GTR, relative risk 0.58 (95% CI: 0.38–0.88, chi-square for heterogeneity 5.72 (3 df), $P=0.13$). The number of sites needed to treat

(NNT) for GTR to achieve buone extra site gaining 2 mm or more attachment over open flap debridement was therefore 8 (95% CI: 4–33), based on the mean incidence of 32% of sites in the control group failing to gain 2 mm or more of attachment. For estimated baseline incidences in the range of the control groups of 10% and 55%, the NNTs are 24 and 3.

Many patients with periodontitis will have only a small number of sites with active disease demonstrating disease progression over the study period and hence only a small number of sites that may be responsive to treatment. In such patient populations, rates of disease progression and mean changes in measures such as probing depth and clinical attachment level over the treatment period are very low. NNT is based on the *difference* in progression rates between the treatment and control arms. In patient populations with low progression rates, differences in progression rates between the treatment and control arms will be small, and the NNT will necessarily be large. In patient populations where the treatment reduces disease-progression rates relative to control, the minimum value of NNT is the inverse of the progression rate in the control group. For example, if the progression rate in the control group is 5%,

the NNT must be at least 20. If the progression rate is low in the control group and is even lower in the treatment arm, the NNT will necessarily be large (Stoner and Payne 2008). As an example, consider a study by Caton et al. (2000), who compared the use of sub-antimicrobial dose doxycycline (SDD) in adult (chronic) periodontitis as an adjunct to scaling and root planing (SRP) with placebo plus SRP. Study endpoints included progression of periodontitis (defined as ≥ 2 mm loss of clinical attachment) over a 9-month treatment period. Among sites with an initial probing depth of at least 7 mm, the reported risk of attachment loss ≥ 2 mm was 0.3% for the SDD plus SRP group and 3.6% for the placebo plus SRP group. The risk difference is 3.3%, which results in a number of sites needed to treat of 31, after rounding up. Therefore, 31 sites on average would need to be treated with the combination of SDD and SRP to avoid periodontitis progression at one additional site relative to treatment with SRP plus placebo (Stoner and Payne 2008).

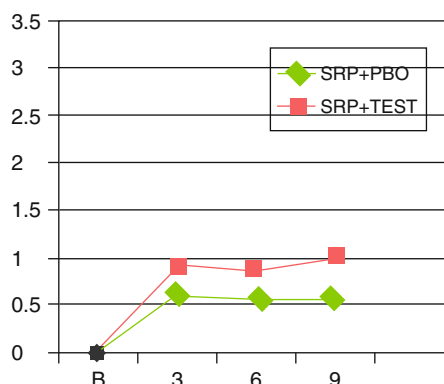


Fig. 9.2 Hypothetical data from a study designed to test the efficacy of scaling and root planing plus chemotherapeutic agents (Jeffcoat 2002). Reprinted with permission from John Wiley & Sons, Inc.)

continue to degenerate over time. The long-term effect of scaling and root planing is affected by patient and clinician compliance with plaque control and maintenance regimens (Jeffcoat 2002).

9.5 Clinical Significance of Periodontal Therapies

9.5.1 Therapies to Reverse Probing Depth or Control the Progression of Periodontitis

Sub-gingival debridement and scaling and root planing are the traditional methods of controlling sub-gingival microflora. The objectives of sub-gingival debridement are to remove not only adherent and unattached bacterial plaque but also, to a lesser extent, deposits of calculus. The primary objective of scaling and root planing is the removal of both calculus and contaminated cementum (Cobb 2002). The effectiveness of either procedure decreases with increasing probing depth, especially when probing depths exceed 5 mm (Cobb 1996). The hypothesis underlying this type of study is that therapy such as scaling and root planing will exhibit some (albeit small) improvements in soft tissue parameters (such as probing depths, clinical attachment levels, bleeding on probing, etc.), and that the control group will con-

9.5.2 Scaling and Root Planing Plus Chemotherapeutic Agents

The underlying hypothesis for these studies is that scaling and root planing plus placebo will exhibit some improvement in soft tissue parameters, and scaling and root planing plus active agent will show more reduction in probing depth, etc. (Fig. 9.2). For agents currently available, this usually occurs over the first year. Part of the determination of clinical significance for most clinicians involves the length of effect. It is unreasonable to expect the currently available drugs to last a lifetime; plaque control and maintenance are also needed. Therefore, clinical significance should be assessed early on. Later assessments need to take plaque control and frequency of maintenance into account. Other clinical issues include the amount of improvement sufficient to justify the time and expense of the use of the new agent. The particular patient population may be especially pertinent here (Jeffcoat 2002).

Systemic medications are of value as adjuncts to periodontal therapy. These medications can be divided into two major categories: antibiotics and

agents for host modulation. Antibiotics have been shown to be valuable adjuncts in specialized types of periodontal disease, such as localized and generalized aggressive periodontitis, and of possible value in severe chronic periodontitis (Ciancio 2002). Local delivery of an antimicrobial offers the clinician a statistically and clinically significant option in the treatment of chronic periodontitis. Greenstein (2003b) investigated the benefits of using minocycline microspheres as an adjunct to conventional periodontal therapy. The author reviewed data from a large controlled clinical trial. The investigation indicated that scaling and root planing, plus minocycline microspheres, attained statistically significant improvements, with regard to several clinical parameters, when compared with scaling and root planing alone. Because it is the clinical meaningfulness of data that determines whether a therapy should be implemented, these data were interpreted with respect to their clinical relevance. Killoy (2002) reviewed what information or data would be helpful in determining clinical significance. It then addressed nine criteria for clinical significance for as many products as possible.

1. Efficacy and statistical significance of all systems are similar or better than the same for periodontal surgery versus scaling and root planing (SRP).
2. Chlorhexidine chip, doxycycline gel, and minocycline microspheres as adjunctive therapy showed a significant improvement of 2 mm or more in PD that was also significantly better than the control SRP. Doxycycline gel as an adjunctive therapy did the same for CAL.
3. Doxycycline gel and minocycline microspheres as an adjunct showed a low and non-significant worsening of 2 mm or more in PD. Doxycycline gel as an adjunct did similarly for CAL.
4. Doxycycline gel with debridement in pockets initially 5–6 mm showed an improvement to ≤ 5 mm of 69% of the sites at 3 months. These sites did not require additional therapy. This was significantly better than the SRP alone (60%).
5. Doxycycline gel with debridement in all pockets initially > 5 mm showed an improvement to ≤ 5 mm of 58% of the sites at 3 months. These sites did not require additional therapy. This was significantly better than SRP alone (50%).
6. Minocycline microspheres with SRP in pockets initially > 6 mm reduced PD to < 5 mm at 9 months in 65% of the sites, which was significantly better than SRP (35%).
7. Morbidity for all systems is low and similar to SRP.
8. Chlorhexidine chip, doxycycline gel, and minocycline microspheres require considerably less time to place than the time taken to perform SRP.
9. Combined use of doxycycline gel and debridement significantly reduced total treatment time compared with SRP (1 h and 11 min).
10. The cost of adding chlorhexidine chip to treatment regimens is partially offset by reduced surgical costs.

9.5.3 Regenerative Therapy

Although regenerative therapy is not part of this symposium, the study hypothesis is that the vehicle control may exhibit some regeneration, but the test agent will exhibit more regeneration (Fig. 9.3). Usually, regeneration is detectable over the first year, and the duration of effect is not well established. The amount of regeneration should be sufficient to justify the time and expense of surgery (Jeffcoat 2002).

In summary, the following was proposed by Addy and Newcombe (2005) to assess the clinical significance of prevention or therapeutic regimens in periodontics:

- Benchmark efficacy: achieving an outcome similar to or better than other established regimens
- Positive efficacy: achieving an effect similar to or greater than the most effective regimen to date
- Disease efficacy: achieving an effect on an etiological factor which reduces a disease (plaque to gingivitis or plaque to primary or secondary periodontitis)
- Proportional efficacy: achieving a previously agreed proportional reduction in a parameter compared with control

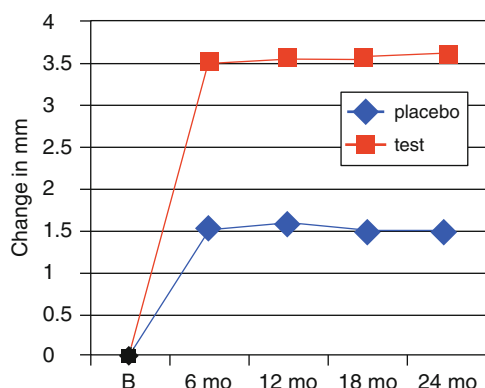


Fig. 9.3 Hypothetical data from a study designed to test the efficacy of regenerative therapy (Jeffcoat 2002. Reprinted with permission from John Wiley & Sons, Inc.)

Where the outcome measure is dental plaque or gingivitis, the magnitude of the effect over control for clinical significance is likely to be high except in the case of positive efficacy. Where the outcome measure is the response of periodontitis to treatment (probing depth, loss of attachment) the magnitude of the difference from the benchmark or positive control for clinical significance is likely to be small in both absolute (millimetres) and proportionate terms. The decision concerning clinical significance may be modulated by secondary factors, in particular, cost/benefit ratio, side effects, patient comfort, and risk (Addy and Newcombe 2005).

References

- Addy M, Newcombe RG. Statistical versus clinical significance in periodontal research and practice. *Periodontol* 2000. 2005;39:132–44.
- Caton JG, Ciancio SG, Blieden TM, Bradshaw M, Crout RJ, Hefti AF, Massaro JM, Polson AM, Thomas J, Walker C. Treatment with subantimicrobial dose doxycycline improves the efficacy of scaling and root planing in patients with adult periodontitis. *J Periodontol*. 2000;71(4):521–32.
- Chatellier G, Zapletal E, Lemaitre D, Menard J, Degoulet P. The number needed to treat: a clinically useful nomogram in its proper context. *BMJ*. 1996;312(7028):426–9.
- Christley RM. Statistical significance, power and sample size – what does it all mean? *J Small Anim Pract*. 2008a;49(6):263.
- Christley RM. Statistical significance, power and sample size—part 2. *J Small Anim Pract*. 2008b;49:321–2.
- Ciancio SG. Systemic medications: clinical significance in periodontics. *J Clin Periodontol*. 2002;29 Suppl 2:17–21.
- Cobb CM. Non-surgical pocket therapy: mechanical. *Ann Periodontol*. 1996;1:443–90.
- Cobb CM. Clinical significance of non-surgical periodontal therapy: an evidence-based perspective of scaling and root planing. *J Clin Periodontol*. 2002;29:6–16.
- Cohen J. Statistical power analysis for the behavioural sciences. 2nd ed. Hillsdale: Erlbaum; 1988.
- Godwin M. Hypothesis: the research page. Part 3: power, sample size, and clinical significance. *Can Fam Physician*. 2001;47(1441–3):1450–3.
- Greenstein G. Clinical versus statistical significance as they relate to the efficacy of periodontal therapy. *J Am Dent Assoc*. 2003a;134:583–91.
- Greenstein G. The efficacy of minocycline microspheres in the treatment of chronic periodontitis: statistical vs clinical significance. *Compend Contin Educ Dent*. 2003b;24:121–6, 128, 130.
- Greenstein G. Local drug delivery in the treatment of periodontal diseases: assessing the clinical significance of the results. *J Periodontol*. 2006;77:565–78.
- Greenstein G, Lamster I. Efficacy of periodontal therapy: statistical versus clinical significance. *J Periodontol*. 2000;71:657–62.
- Greenstein G, Nunn ME. A method to enhance determining the clinical relevance of periodontal research data: number needed to treat (NNT). *J Periodontol*. 2004;75:620–4.
- Hollon SD, Flick SN. On the meaning if clinical significance. *Behav Assess*. 1988;10:197–206.
- Hujoel PP. Endpoints in periodontal trials: the need for an evidence-based research approach. *Periodontol* 2000. 2004;36:196–204.
- Hujoel PP, Leroux BG, DeRouen TA, Powell LV, Kiyak HA. Evaluating the validity of probing attachment loss as a surrogate for tooth mortality in a clinical trial on the elderly. *J Dent Res*. 1997;76:858–66.
- Hujoel PP, Löe H, Anerud A, Boysen H, Leroux BG. The informativeness of attachment loss on tooth mortality. *J Periodontol*. 1999;70:44–8.
- Hujoel PP, Armitage GC, Garcia RI. A perspective on clinical significance (editorial). *J Periodontol*. 2000;71:1515–8.
- Hutton JL. Number needed to treat and number needed to harm are not the best way to report and assess the results of randomised clinical trials. *Br J Haematol*. 2009;146:27–30.
- Jeffcoat M. What is clinical significance? *J Clin Periodontol*. 2002;29 Suppl 2:30–2.
- Killooy WJ. The clinical significance of local chemotherapies. *J Clin Periodontol*. 2002;29 suppl 2:22–9.
- Kingman A. Statistical vs clinical significance in product testing: can they be designed to satisfy equivalence? *J Public Health Dent*. 1992;52:353–60.
- LeFort SM. The statistical versus clinical significance debate. *Image J Nurs Sch*. 1993;25:57–62.

- Lindgren BR, Wielinskyi CL, Finkelstein SM, Warwick WJ. Contrasting clinical and statistical significance within the research setting. *Pediatr Pulmonol*. 1993;16:336–40.
- McQuay HJ, Moore RA. Using numerical results for systematic reviews in clinical practice. *Ann Intern Med*. 1997;126:712–20.
- Needleman I, Moles DR, Worthington H. Evidence-based periodontology, systematic reviews and research quality. *Periodontol* 2000. 2005;37:12–28.
- Needleman I, Tucker R, Giedrys-Leeper E, Worthington H. A systematic review of guided tissue regeneration for periodontal infrabony defects. *J Periodontol Res*. 2002;37(5):380–8.
- Rethman MP, Nunn ME. Clinical versus statistical significance (editorial). *J Periodontol*. 1999;70:700–2.
- Singh G. A shift from significance test to hypothesis test through power analysis in medical research. *J Postgrad Med*. 2006;52:148–50.
- Stoner JA, Payne JB. Interpretation of treatment effects in periodontal research: a note on the number needed to treat. *J Can Dent Assoc*. 2008;74(5):435–7.
- Tramèr MR, Walder B. Number needed to treat (or harm). *World J Surg*. 2005;29:576–81.
- Van Dyke TE. The clinical significance of new therapies for the management of periodontal disease. *J Int Acad Periodontol*. 2005;7:191–6.

Clinical practice guidelines are systematically developed statements that aim to help physicians and patients reach the best health-care decisions. Good guidelines have many attributes, including validity, reliability, reproducibility, clinical applicability and flexibility, clarity, development through a multidisciplinary process, scheduled reviews, and documentation. More than 2,500 guidelines are currently represented in the National Guideline Clearinghouse (www.guideline.gov). Medical specialty societies are their most common sponsors (Steinbrook 2007).

The broad interest in clinical guidelines that is stretching across Europe, North America, Australia, New Zealand, and Africa has its origin in issues that most health-care systems face: rising health-care costs, fuelled by increased demand for care, more expensive technologies, and an ageing population; variations in service delivery among providers, hospitals, and geographical regions and the presumption that at least some of this variation stems from inappropriate care, either overuse or underuse of services; and the intrinsic desire of health-care professionals to offer, and of patients to receive, the best care possible. Clinicians, policymakers, and payers see guidelines as a tool for making care more consistent and efficient and for closing the gap between what clinicians do and what scientific evidence supports (Woolf et al. 1999).

Among the efforts in the United States that are generally considered successful in developing high-quality clinical guidelines are those of the US Preventive Services Task Force, the Advisory

Committee on Immunization Practices, and the National Academies, as well as the treatment guidelines issued by the Centers for Disease Control and Prevention. Efforts outside the United States include those of the World Health Organization and the National Institute for Clinical Excellence (NICE), in the United Kingdom (Steinbrook 2007).

10.1 Benefits of Clinical Guidelines

The principal benefit of guidelines is to improve the quality of care received by patients. Although it has been shown in rigorous evaluations that clinical practice guidelines can improve the quality of care, whether they achieve this in daily practice is less clear. This is partly because patients, doctors, payers, and managers define quality differently and because current evidence about the effectiveness of guidelines is incomplete (Woolf et al. 1999).

10.1.1 Potential Benefits for Patients

For patients (and almost everyone else in health care), the greatest benefit that could be achieved by guidelines is to improve health outcomes. Guidelines that promote interventions of proved benefit and discourage ineffective ones have the potential to reduce morbidity and mortality and improve quality of life, at least for some conditions. Guidelines can also improve the consistency

of care; patients' studies around the world with identical clinical problems receive different care depending on their clinician, hospital, or location. Guidelines offer a remedy, making it more likely that patients will be cared for in the same manner regardless of where or by whom they are treated (Woolf et al. 1999).

Clinical guidelines offer patients other benefits. Those accompanied by 'consumer' versions (leaflets, audiotapes, or videos in lay language) or publicized in magazines, news reports, and Internet sites inform patients and the public about what their clinicians should be doing. Some guidelines empower patients to make more informed health-care choices and to consider their personal needs and preferences in selecting the best option. Indeed, clinicians may first learn about new guidelines (or be reminded of oversights) when patients ask about recommendations or treatment options (Woolf et al. 1999).

Finally, clinical guidelines can help patients by influencing public policy. Guidelines call attention to underrecognized health problems, clinical services, and preventive interventions and to neglected patient populations and high-risk groups. Services that were not previously offered to patients may be made available as a response to newly released guidelines. Clinical guidelines developed with attention to the public good can promote distributive justice, advocating better delivery of services to those in need. In a cash-limited health-care system, guidelines that improve the efficiency of health care free up resources needed for other (more equitably distributed) health-care services (Woolf et al. 1999).

10.1.2 Potential Benefits for Health-Care Professionals

Clinical guidelines can improve the quality of clinical decisions. They offer explicit recommendations for clinicians who are uncertain about how to proceed, overturn the beliefs of doctors accustomed to outdated practices, improve the consistency of care, and provide authoritative

recommendations that reassure practitioners about the appropriateness of their treatment policies. Guidelines based on a critical appraisal of scientific evidence (evidence-based guidelines) clarify which interventions are of proved benefit and document the quality of the supporting data. They alert clinicians to interventions unsupported by good science, reinforce the importance and methods of critical appraisal, and call attention to ineffective, dangerous, and wasteful practices (Woolf et al. 1999).

10.1.3 Potential Benefits for Health-Care Systems

Health-care systems that provide services, and government bodies and private insurers that pay for them, have found that clinical guidelines may be effective in improving efficiency (often by standardizing care) and optimizing value for money. Implementation of certain guidelines reduces outlays for hospitalization, prescription drugs, surgery, and other procedures. Publicizing adherence to guidelines may also improve public image, sending messages of commitment to excellence and quality. Such messages can promote good will, political support, and (in some health-care systems) revenue. Many believe that the economic motive behind clinical guidelines is the principal reason for their popularity (Woolf et al. 1999).

10.2 Limits of Clinical Guidelines

Guidelines may be controversial for many reasons (Steinbrook 2007):

1. For recommending too little and for recommending too much.
2. For contradicting dogma or for providing reasons for insurers to deny coverage for specific drugs or devices.
3. For financial conflict, when pharmaceutical and medical device companies with a financial stake in the outcome provide substantial funding for their development and implementation.

When members of guideline committees also have substantial financial associations with industry, further questions inevitably arise (Steinbrook 2007).

Examples of financial conflicts of interests include ownership of stocks or shares, paid employment or consultancy, board membership, patent applications, research grants (from any source, restricted or unrestricted), travel grants and honoraria for speaking or participating at meetings, and gifts. Examples of non-financial conflicts of interests include leadership or close involvement in an advocacy group that stands to gain from a Clinical Guidelines Committee member's opinion; being a chair or member of another guideline committee relevant to the topic under discussion; acting as an expert witness or having a membership (in a government or other advisory board) or relationship (paid or unpaid) with organizations and funding bodies (including non-governmental organizations, research institutions, or charities), or a membership in a lobbying or advocacy organization; writing or consulting for an educational company; having personal relationships (i.e. a friend, spouse, family member, current or previous mentor, or adversary) with persons involved in the submission or evaluation of a paper, such as authors, reviewers, editors, or members of the editorial board of a *Public Library of Science* journal; and having personal convictions (political, religious, ideological, or other) related to a paper's topic that may interfere with an unbiased publication process (at the stage of authorship, peer review, editorial decision-making, or publication) (Qaseem et al. 2010).

10.2.1 Potential Harms to Patients

The greatest danger of flawed clinical guidelines is to patients. Recommendations that do not take due account of the evidence can result in suboptimal, ineffective, or harmful practices. Guidelines that are inflexible can harm by leaving insufficient room for clinicians to tailor care to patients' personal circumstances and medical history. Recommendations against an intervention

may lead providers to drop access to or coverage for services. Imprudent recommendations for costly interventions may displace limited resources that are needed for other services of greater value to patients. The tendency of guidelines to focus attention on specific health issues is subject to misuse by proponents and advocacy groups, giving the public (and health professionals) the wrong impression about the relative importance of diseases and the effectiveness of interventions (Woolf et al. 1999).

10.2.2 Potential Harms to Health-Care Professionals

Flawed clinical guidelines harm practitioners by providing inaccurate scientific information and clinical advice, thereby compromising the quality of care. They may encourage ineffective, harmful, or wasteful interventions. Even when guidelines are correct, clinicians often find them inconvenient and time consuming to use. Outdated recommendations may perpetuate outmoded practices and technologies.

Guidelines can harm medical investigators and scientific progress if further research is inappropriately discouraged. Guidelines that conclude that a procedure or treatment lacks evidence of benefit may be misinterpreted by funding bodies as grounds for not investing in further research and for not supporting efforts to refine previously ineffective technologies (Woolf et al. 1999).

10.2.3 Potential Harms to Health-Care Systems

Health-care systems and payers may be harmed by guidelines if following them escalates utilization, compromises operating efficiency, or wastes limited resources. Some clinical guidelines, especially those developed by medical and other groups unconcerned about financing, may advocate costly interventions that are unaffordable or that cut into resources needed for more effective services (Woolf et al. 1999).

10.3 Development of Clinical Practice Guidelines and Clinical Guidance Statements

10.3.1 Proposed Guiding Principles Underlying the Guideline Development Process

National Health and Medical Research Council Australia (1998) describes in detail the nine guiding principles underlying the guideline development process. They are as follows:

1. **Processes for guideline development and evaluation should be outcome focused.** The process for guideline development should be aimed at identifying interventions that will ensure the best possible health outcomes. To date, much of the evaluation of health care has centred around the process of health care (whether clinicians conform to recommended practices) rather than the outcomes (whether the recommended practices produce a change in health).
2. **Guidelines should be based on the best available evidence and should include a statement about the strength of recommendations.** Ideally, recommendations should be based on the highest level of evidence, preferably a systematic review of high-quality randomized controlled clinical trials that measure relevant outcomes and demonstrate a strong, clinically important, beneficial effect of the intervention.
3. **The method used to synthesize the available evidence should be the strongest applicable.**
4. **The process of guideline development should be multidisciplinary and should include consumers.** The involvement of a range of specialist and generalist clinicians, allied health professionals, experts in methodology (such as epidemiologists and health economists), and consumers will improve the quality and continuity of care and increase the likelihood of the guidelines' adoption.
5. **Guidelines should be flexible and capable of adapting to varying local conditions.** Guidelines should provide a clear statement of the evidence but allow for flexibility and adaptability in implementation. Thus, they should

provide evidence relevant to different target populations, which may vary in terms of age, gender, ethnicity, diagnosis, disease severity, co-morbidity, and social support; take into account resource benefits, costs, and constraints; and offer a means of accommodating different consumer values and preferences.

6. **Guidelines should be developed with resource constraints in mind.**
7. **Guidelines should be developed, disseminated, and implemented taking into account their target audiences.** It is important that an economic appraisal be incorporated in guidelines. In many circumstances, this may be useful in guiding clinical decisions about treatment options. There are three types of economic appraisal comparing two interventions, or one intervention against a placebo:
 - A cost-minimization analysis
 - A cost-effectiveness analysis
 - A cost-utility analysis
8. **The validity and usefulness of the guidelines should be evaluated.** It is important to evaluate the implementation process to determine the extent to which the guidelines affected practitioners' knowledge and behaviour and what, if any, factors contributed to non-compliance with the guidelines. These results might inform the development of more effective implementation strategies
9. **Guidelines should be revised regularly.** Because guidelines need to be based on the best available evidence, they should be reviewed regularly and modified where necessary to take into account new research, new technologies, and the results of evaluation of guideline outcomes.

10.3.2 Current Methods for Developing Practice Guidelines

Current methods for developing practice guidelines include (Woolf 1992):

1. Informal consensus development is the oldest and most common approach, but guidelines produced in this manner are often of poor quality and lack adequate documentation of methods.

2. Formal consensus development uses a systematic approach to assess expert opinion and to reach agreement on recommendations.
3. Evidence-based guideline development links recommendations directly to scientific evidence of effectiveness; rules of evidence are emphasized over expert opinion in making recommendations.
4. Explicit guideline development clarifies the rationale by specifying the potential benefits, harms, and costs of available interventions; estimating the possibility of the outcomes; and comparing the desirability of the outcomes based on patient preferences.

10.3.3 Steps in the Development of Practice Guidelines

The World Health Organization (WHO), like many other organizations around the world, has recognized the need to use more rigorous processes to ensure that health-care recommendations are informed by the best available research evidence. Guidelines for Guidelines (2006) (<http://hiru.mcmaster.ca/cochrane/cochrane/hbook.htm>) developed as a WHO series identified 19 components that should be included in the 'Guidelines for WHO Guidelines' (GWG) (Schünemann et al. 2006a). The components are:

10.3.3.1 Priority Setting

The following criteria for establishing priorities for developing recommendations based on the WHO's aims and strategic advantages were suggested (Oxman et al. 2006a):

- Problems associated with a high burden of illness in low- and middle-income countries, or new and emerging diseases.
- No existing recommendations of good quality.
- The feasibility of developing recommendations that will improve health outcomes, reduce inequities, or reduce unnecessary costs if they are implemented.
- Implementation is feasible, will not exhaustively use available resources, and barriers to change are not likely to be so high that they cannot be overcome.

- Additional priorities for the WHO include interventions that will likely require system changes and interventions where there might be a conflict in choices between individual and societal perspectives (Oxman et al. 2006a).

10.3.3.2 Group Composition (and Consultations)

The existing empirical evidence suggests that panel composition has an impact on the content of the recommendations that are made. There is limited research evidence to guide the exact composition of a panel. Based on logical arguments and the experience of other organizations, it was recommended the following:

- Groups that develop guidelines or recommendations should be broadly composed and include important stakeholders such as consumers, health professionals that work within the relevant area, and managers or policymakers.
- Groups should include or have access to individuals with the necessary technical skills, including information retrieval, systematic reviewing, health economics, group facilitation, project management, writing, and editing.
- Groups should include or have access to content experts.
- To work well, a group needs an effective leader, capable of guiding the group in terms of the task and process, and capable of facilitating collaboration and balanced contribution from all of the group members.
- Because many group members will not be familiar with the methods and processes that are used in developing recommendations, groups should be offered training and support to help ensure understanding and facilitate active participation (Fretheim et al. 2006a).

10.3.3.3 Declaration and Avoidance of Conflicts of Interest

Although there are no empirical studies of the enforcement of conflict of interest policies, descriptive studies of other organizations and institutions suggest that the WHO convened a standing committee to review all financial disclosure statements prior to the commencement of committee meetings/hearings and to make

management recommendations when necessary. A standard policy requiring all financial ties to be made public (i.e. recorded into the meeting minutes) should reduce the number of problematic cases. In instances where the conflicts seem intractable, a recommendation of recusal may be necessary to protect the greater interests of the WHO and its constituents (Boyd and Bero 2006, P4).

10.3.3.4 Group Processes

Various strategies can be adopted to ensure that the group processes in play when panels are developing recommendations are inclusive, so that all voices can be heard and all arguments given fair weight, including (Fretheim et al. 2006c):

- The use of formal consensus development methods, such as the Nominal Group Technique or the Delphi method
- The selection of a group leader who is qualified and responsible for facilitating an appropriate group process

10.3.3.5 Identification of Important Outcomes Including Cost

Methods used by the WHO to identify important outcomes should be:

- Methods of outcome identification should be transparent and explicit.
- The consultation process should start with identification of all relevant outcomes associated with an intervention.
- Those affected, including consumers, should be involved in the selection of outcomes.
- A question-driven approach (what is important?) is preferable to a data-driven approach (what data are at hand?) to identify important outcomes (Schünemann et al. 2006b).

Type of outcomes considered by the WHO should be:

- Desirable (benefits, less burden, and savings) and undesirable effects should be considered in all guidelines.
- Undesirable effects include harms (including the possibility of unanticipated adverse effects), greater burden (e.g. having to go to the doctor), and costs (including opportunity costs).

- Important outcomes (e.g. mortality, morbidity, quality of life) should be preferred over surrogate, indirect outcomes (e.g. cholesterol levels, lung function) that may or may not correlate with patient important outcomes.
- Ethical considerations should be part of the evaluation of important outcomes (e.g. impacts on autonomy).

10.3.3.6 Explicit Definition of the Question and Eligibility Criteria

The types of evidence that should be used to address different types of questions are:

- The most important type of evidence for informing global recommendations is evidence of the effects of the options (interventions or actions) that are considered in a recommendation. This evidence is essential but not sufficient for making recommendations about what to do. Other types of required evidence are largely context specific.
- The study designs to be included in a review should be dictated by the interventions and outcomes being considered. A decision about how broad a range of study designs to consider should be made in relationship to the characteristics of the interventions being considered, what evidence is available, and the time and resources available.
- There is uncertainty regarding what study designs to include for some specific types of questions, particularly for questions regarding population interventions, harmful effects, and interventions where there is only limited human evidence.
- Decisions about the range of study designs to include should be made explicitly.
- Great caution should be taken to avoid confusing a lack of evidence with evidence of no effect and to acknowledge uncertainty.
- Expert opinion is not a type of study design and should not be used as evidence. The evidence (experience or observations) that is the basis of expert opinions should be identified and appraised in a systematic and transparent way (Oxman et al. 2006b).

10.3.3.7 Type of Study Designs for Different Types of Questions

While there is broad agreement that the study designs to be included in a review should be dictated by the interventions being reviewed, there is uncertainty regarding what study designs to include for some specific types of questions. For any question, as the range of study designs that are included is broadened, an increasing amount of work is required to derive decreasingly reliable estimates of the effects of interventions. A decision about how broad a range of study designs to consider must be made in relationship to the characteristics of the interventions, what evidence is available, and the time and resources available. It is important to recognize that decisions about what study designs to include may also be influenced by the extent to which relevant studies are available that have used study designs that are most likely to provide valid results. That is, there may sometimes be a trade-off between including studies that are more likely to be valid and ones that are more likely to be directly relevant (Oxman et al. 2006b).

10.3.3.8 Identification of Evidence

The Guidelines for WHO Guidelines (GWG) state: “*It is recommended that the systematic review be undertaken.*” (<http://hiru.mcmaster.ca/cochrane/cochrane/hbook.htm>). After the studies have been identified and critically appraised, and the evidence synthesized, evidence should be graded. All evidence, including that on safety, should be clearly laid out in an evidence table. Meta-analysis should be done when the data permit (Oxman et al. 2006b).

10.3.3.9 Synthesis and Presentation of Evidence

- Because preparing systematic reviews can take over a year and require capacity and resources, existing reviews should be used when possible and updated, if needed.
- Standard criteria, such as A MeaSurement Tool to Assess Reviews (AMSTAR), should be used to critically appraise existing systematic reviews, together with an assessment of the relevance of the review to the questions being asked.

- Consideration should be given to undertaking or commissioning a new review whenever a relevant, up-to-date review of good quality is not available.
- When time or resources are limited, it may be necessary to undertake rapid assessments. The methods that are used to do these assessments should be reported, including important limitations and uncertainties and explicit consideration of the need and urgency of undertaking a full systematic review.
- Because the WHO has limited capacity for undertaking systematic reviews, reviews will often need to be commissioned when a new review is needed. Consideration should be given to establishing collaborating centres to undertake or support this work similar to what some national organizations have done (Oxman et al. 2006c).

10.3.3.10 Specification and Integration of Values

- Recommendations should include a description of how decisions were made about the relative importance of the consequences (benefits, harms, and costs) of a decision.
- Values that influence recommendations should be reported along with the research evidence underlying recommendations.
- When differences in values would lead to different decisions or there is important uncertainty about values that are critical to a decision, this should be flagged and reflected in the strength of the recommendation.
- Adaptable guideline templates that allow for integration of different values should be developed and used when differences in values are likely to be critical to a decision (Schünemann et al. 2006d).

10.3.3.11 Making Judgments About Desirable and Undesirable Effects

Reports should be structured using headings that correspond to those suggested by the Conference on Guideline Standardization or similar headings. The type and magnitude of the main benefits, harms, and costs that are expected to result from

guideline implementation should be reported (Schünemann et al. 2006c, d; Oxman et al. 2006e).

10.3.3.12 Taking Account of Equity

The following question should routinely be considered when inequities are addressed in systematic reviews that are used as background documents for recommendations: Are there plausible reasons for anticipating differential relative effects across disadvantaged and advantaged populations? If there are plausible reasons for anticipating differential effects, additional evidence should be included in a review to inform judgments about the likelihood of differential effects (Oxman et al. 2006d).

The following additional questions should routinely be considered by those making recommendations on behalf of the WHO (Oxman et al. 2006d):

- How likely is it that the results of available research are applicable to disadvantaged populations and settings?
- How likely are differences in baseline risk that would result in differential absolute effects across disadvantaged and advantaged populations?
- How likely is it that there are important differences in trade-offs between the expected benefits and harms across disadvantaged and advantaged populations?
- Are there different implications for disadvantaged and advantaged populations, or implications for addressing inequities?

10.3.3.13 Grading Evidence and Recommendations

- Both the quality of evidence and the strength of recommendations should be graded. The criteria used to grade the strength of recommendations should include the quality of the underlying evidence but should not be limited to that.
- The approach to grading should be one that has wide international support and is suitable for a wide range of different types of recommendations. The **Grading of Recommendations Assessment, Development and Evaluation (GRADE)** approach, which is currently sug-

gested in the Guidelines for WHO Guidelines, is being used by an increasing number of other organizations internationally. It should be used more consistently by the WHO. Further developments of this approach should ensure its wide applicability (Schünemann et al. 2006c).

While the grading of recommendations represents a positive development for guideline development and interpretation, the proliferation of grading systems has proved to be an unfortunate consequence. Methodologists and guideline developers have given much thought and effort to considering the criteria and approaches to an optimal grading system. The American College of Chest Physicians (ACCP) convened a working group to review the issue and to agree on a grading system that would be consistent with the latest developments in the field. The task force began by developing criteria that define an optimal grading system, placing them in an order that approximates their relative importance. These criteria guided the decisions of the group in the choice of the grading system that follows (Guyatt et al. 2006).

The Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) working group is a group of health professionals, researchers, and guideline developers worldwide who, in 2000, began to work together to develop an optimal system of rating quality of evidence and determining strength of recommendations for clinical practice guidelines (Guyatt et al. 2011a, b).

The published literature includes a number of articles describing the GRADE approach, of which the most comprehensive is a six-part series published in 2008 in the *BMJ* (Guyatt et al. 2008a, b, c, d; Jaeschke et al. 2008). The audience for these articles is, however, the clinician and policy-making users of GRADE's output, which includes evidence profiles, summary of findings tables, and graded recommendations (all facilitated by a computer program, GRADEpro, that the working group has produced) (Guyatt et al. 2011a, b).

What previous articles fail to do is provide detailed guidance for those responsible for using GRADE to produce this output: systematic review and health technology assessment authors and the guideline panellists and methodologists who provide support for guideline panels. A

Table 10.1 GRADE Journal of Clinical Epidemiology series list of articles (Guyatt et al. 2011a) (Reprinted with permission from Elsevier)

Introductory articles	1. Introduction and summary of findings tables 2. Framing the question and deciding on the importance of outcomes 3. Rating the quality of evidence-introduction
Rating the quality of evidence	4. Rating the quality of evidence-risk of bias 5. Rating the quality of evidence-publication bias 6. Rating the quality of evidence-imprecision (random error) 7. Rating the quality of evidence-inconsistency 8. Rating the quality of evidence-indirectness 9. Rating up the quality of evidence 10. Rating the quality of evidence for resource use
Summarizing the evidence	11. Summarizing the quality of evidence for individual outcomes and across outcomes 12. Preparing summary of findings tables-binary outcomes 13. Preparing summary of findings tables-continuous outcomes
Diagnostic tests	14. Applying GRADE to diagnostic tests
Making recommendations	15. Going from evidence to recommendations-the meaning of strong and weak recommendations 16. Going from evidence to recommendations-determinants of a recommendation's direction and strength 17. Going from evidence to recommendations-resource use
GRADE and observational studies	18. Special challenges in using observational studies
Concluding articles	19. Group processes, variations of GRADE, and further developments of GRADE part 1 20. Group processes, variations of GRADE, and further developments of GRADE part 2

series of articles published in *Journal of Clinical Epidemiology* address this deficiency. This series, which provides guidance for each step in the application of GRADE, includes 20 articles (Table 10.1) (Guyatt et al. 2011a, b).

Figure 10.1 presents a schematic view of GRADE's process for developing recommendations in which unshaded boxes describe steps in the process common to systematic reviews and guidelines and the shaded boxes describe steps that are specific to guidelines. One begins by defining the question in terms of the populations, alternative management strategies (an intervention, sometimes experimental and a comparator, sometimes standard care), and all patient important outcomes (in this case four). For guidelines, one classifies those outcomes as either critical (two outcomes in the figure) or important but not critical (two outcomes). A systematic search leads to inclusion of relevant studies (in this schematized presentation, five such studies) (Guyatt et al. 2011b).

Although the quality of evidence represents a continuum, the GRADE approach results in an assessment of the quality of a body of evidence as high, moderate, low, or very low. Table 10.2 presents what GRADE means by each of these four categories and contrasts their current definition with the previous definition, which focused on the implications of the levels of evidence for future research (the lower the quality, the more likely further research would change our confidence in the estimates, and the estimates themselves). The earlier characterization has been criticized because there are many situations in which we cannot expect higher quality evidence to be forthcoming (Balslem et al. 2011).

Table 10.3 summarizes GRADE's approach to rating the quality of evidence, which begins with the study design (trials or observational studies) and then addresses five reasons to possibly rate down the quality of evidence and three to possibly rate up the quality (Balslem et al. 2011).

Fig. 10.1 Schematic view of GRADE’s process for developing recommendations. *Abbreviation: RCT* randomized controlled trials. *Also labelled “conditional” or “discretionary” (Guyatt et al. (2011b). Reprinted with permission from Elsevier)

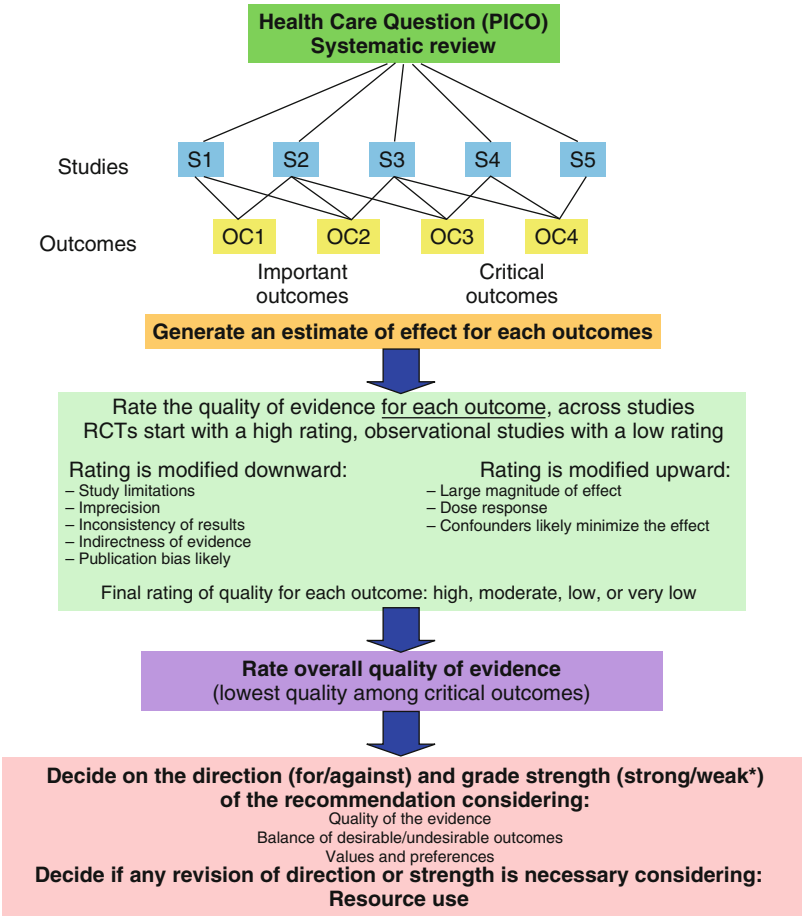


Table 10.2 Significance of the four levels of evidence (Balshem et al. 2011) (Reprinted with permission from Elsevier)

Quality level	Current definition	Previous definition
High	We are very confident that the true effect lies close to that of the estimate of the effect	Further research is very unlikely to change our confidence in the estimate of effect
Moderate	We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect	Any estimate of effect is very uncertain

The GRADE working group has developed specific approaches to presenting the quality of the available evidence, the judgments that bear on the quality rating, and the effects of alternative

management strategies on the outcomes of interest. These approaches are now summarized in the GRADE EP and the SoFs table. An **EP table** (Table 10.4) includes a detailed quality assess

Table 10.3 A summary of GRADE’s approach to rating quality of evidence (Balshem et al. 2011) (Reprinted with permission from Elsevier)

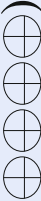
Study design	Initial quality of a body of evidence	Lower if	Higher if	Quality of a body of evidence
Randomized trials	High	Risk of bias –1 Serious –2 Very serious	Large effect +1 Large +2 Very large	High (four plus: )
		Inconsistency –1 Serious –2 Very serious	Dose response +1 Evidence of a gradient	Moderate (three plus: )
Observational studies	Low	Indirectness –1 Serious –2 Very serious	All plausible residual confounding	Low (two plus: )
		Imprecision –1 Serious –2 Very serious Publication bias –1 Likely –2 Very likely	+1 Would reduce a demonstrated effect +1 Would suggest a spurious effect if no effect was observed	Very low (one plus: )

Table 10.4 Example of GRADE EP Table antibiotics for children with acute otitis media (Guyatt et al. 2011b) (Reprinted with permission from Elsevier)

Quality assessment				Summary of findings							
No of studies (design)	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Number of patients		Relative risk (95% CI)	Absolute risk		Quality
						Placebo	Antibiotics		Control risk ^a	Risk difference (95% CI)	
Pain at 24 h 5 (RCT)	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	241/605	223/624	RR 0.9 (0.78–1.04)	367/1,000	Not significant	    High
Pain at 2–7 day 10 (RCT)	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	303/1,366	228/1,425	RR 0.72 (0.62–0.83)	257/1,000	72 fewer per 1,000(44–98)	    High
Hearing, inferred from the surrogate outcome abnormal tympanometry–1 mo											
1 month 4 (RCT)	No serious limitations	No serious inconsistency	Serious indirectness (because of indirectness of outcome)	No serious imprecision	Undetected	168/460	153/467	RR 0.89 (0.75–1.07)	350/1,000	Not significant	    Moderate
Hearing, inferred from the surrogate outcome abnormal tympanometry 3 mo											
3 (RCT)	No serious limitations	No serious inconsistency	Serious indirectness (because of indirectness of outcome)	No serious imprecision	Undetected	96/398	96/410	RR 0.97 (0.76–1.24)	234/1,000	Not significant	    Moderate
Vomiting, diarrhoea, or rash											
5 (RCT) No serious indirectness	No serious limitations	Serious inconsistency (because of inconsistency in absolute effects)	No serious imprecision	Undetected	83/711	110/690	RR 1.38 (1.09–1.76)	113/1,000	43 more per 1,000 (10–86)	Moderate	    Moderate

Abbreviations: GRADE, Grading of Recommendations Assessment, Development, and Evaluation; RCT, randomized controlled trials; CI, confidence interval; RR, risk ratio. ^aThe control rate is based on the median control group risk across studies.

ment in addition to a SoFs. That is, the EP includes an explicit judgment of each factor that determines the quality of evidence for each outcome, in addition to a SoFs for each outcome. The **SoF table** (Table 10.5) includes an assessment of the quality of evidence for each outcome but not the detailed judgments on which that assessment is based (Guyatt et al. 2011b).

The EP and the SoF tables serve different purposes and are intended for different audiences. The EP provides a record of the judgments that were made by review or guideline authors. It is intended for review authors, those preparing SoF tables and anyone who questions a quality assessment. It helps those preparing SoF tables to ensure that the judgments they make are systematic and transparent and it allows others to inspect those judgments. Guideline panels should use EPs to ensure that they agree about the judgments underlying the quality assessments and to establish the judgments recorded in the SoF tables. SoF tables are intended for a broader audience, including end users of systematic reviews and guidelines. They provide a concise summary of the key information that is needed by someone making a decision and, in the context of a guideline, provide a summary of the key information underlying a recommendation. GRADEpro computer software facilitates the process of developing both EPs and SoFs tables (Guyatt et al. 2011b).

The authors of the GRADE system suggest the four major factors may affect a recommendation: (1) balance between desirable and undesirable effects, (2) the quality of the evidence, (3) values and preferences, and (4) cost–resource allocation (Faggion 2010) (Fig. 10.2).

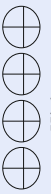
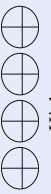
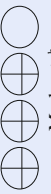
A large number of taxonomies are used to rate the quality of an individual study and the strength of a recommendation based on a body of evidence. Ebell et al. (2004) have developed a new grading scale, called the **Strength of Recommendation Taxonomy (SORT)**, that will be used by several family medicine and primary care journals (required or optional), with the goal of allowing readers to learn one taxonomy that will apply to many sources of evidence. It addresses the quality, quantity, and consistency of evidence and

allows authors to rate individual studies or bodies of evidence.

- *Quality of evidence:* The SORT system furnishes 3 levels of quality (Table 10.6). The classification is clearly divided into patient-oriented evidence (levels 1 and 2) and disease-oriented evidence (level 3). POEMs, or ‘patient-oriented evidence that matters’, allows clinicians to filter information from the medical literature and focus only on what is in fact important for the patient. DOE, or ‘disease-oriented evidence’, includes intermediate, histopathologic, physiologic, or surrogate results (e.g. pocket probing depth changes) that may or may not reflect improvement in patient outcomes. Levels of evidence 1 and 2 are differentiated by the design and quality of the studies (Table 10.6) (Faggion 2010; Ebell et al. 2004).
- *Strength of recommendations:* The taxonomy emphasizes the use of patient-oriented outcomes that measure changes in morbidity or mortality. Disease-oriented evidence will always generate recommendation grade C (the weakest recommendation). The consistency of the body of evidence will determine the choice between grades A and B. The rationale is described in detail in Table 10.6 (Faggion 2010; Ebell et al. 2004).
- *Strengths and weaknesses:* The main criteria for determining the level of evidence and the strength of a recommendation (type of evidence—disease or patient oriented) might facilitate use of this system in clinical dentistry. Currently, most evidence on dental treatments is disease oriented (e.g. changes in probing-pocket depth [PPD], and clinical attachment level [CAL] in periodontal treatments), and determination of evidence level C (and consequently recommendation C) by the clinician is straightforward. This system, however, does not allow patient preference to affect the decision-making process, because it emphasizes internal validity of studies (Faggion 2010; Ebell et al. 2004).

The American College of Physicians’s grading system for the quality of evidence and strength of recommendations was adapted from the

Table 10.5 Example of GRADE SoF Table: summary of finding: antibiotics for acute otitis media in children (Guyatt et al. 2011b) (Reprinted with permission from Elsevier)

Antibiotics compared with placebo for acute otitis media in children							
Patient or population: children with acute otitis media							
Setting: high- and middle-income countries							
Intervention: antibiotics							
Comparison: placebo							
Outcomes	Estimated risks (95% CI)		Relative effect (95% CI)	No. of participants (studies)	Quality of the evidence (GRADE)		Comments
	Control risk ^a	Intervention risk					
Pain at 24 h	Placebo 367 per 1,000	Antibiotics 330 per 1,000 (286–382)	RR 0.9 (0.78–1.04)	1,229 (5)	 High		
Pain at 2–7 day	257 per 1,000	185 per 1,000 (159–213)	RR 0.72 (0.62–0.83)	2,791 (10)	 High		
Hearing, inferred from the surrogate outcome abnormal tympanometry ^d – 1 month	350 per 1,000	311 per 1,000 (262–375)	RR 0.89 (0.75–1.07)	927 (4)	 Moderate ^b		
Hearing, inferred from the surrogate outcome abnormal tympanometry ^d – 3 month	234 per 1,000	227 per 1,000 (178–290)	RR 0.97 (0.76–1.24)	808 (3)	 Moderate ^b		
Vomiting, diarrhoea, or rash	113 per 1,000	156 per 1,000 (123–199)	RR 1.38 (1.09–1.76)	1,401 (5)	 Moderate ^c		Ideally, evidence from nonotitis trials with similar ages and doses (not obtained) might improve the quality of the evidence

CI confidence interval, RR risk ratio, GRADE Grading of Recommendations Assessment, Development, and Evaluation

^aThe basis for the control risk is the median control group risk across studies. The intervention risk (and its 95% CI) is based on the control risk in the comparison group and the relative effect of the intervention (and its 95% CI)

^bBecause of indirectness of outcome

^cGenerally, GRADE rates down for inconsistency in relative effects (which are not inconsistent in this case). Inconsistency here is in absolute effects, which range from 1% to 56%. Contributing factors to the decision to rate down in quality include the likely variation between antibiotics and the fact that most of the adverse events come from a single study. Consideration of indirect evidence from other trials of antibiotics in children (not undertaken) would likely further inform this issue

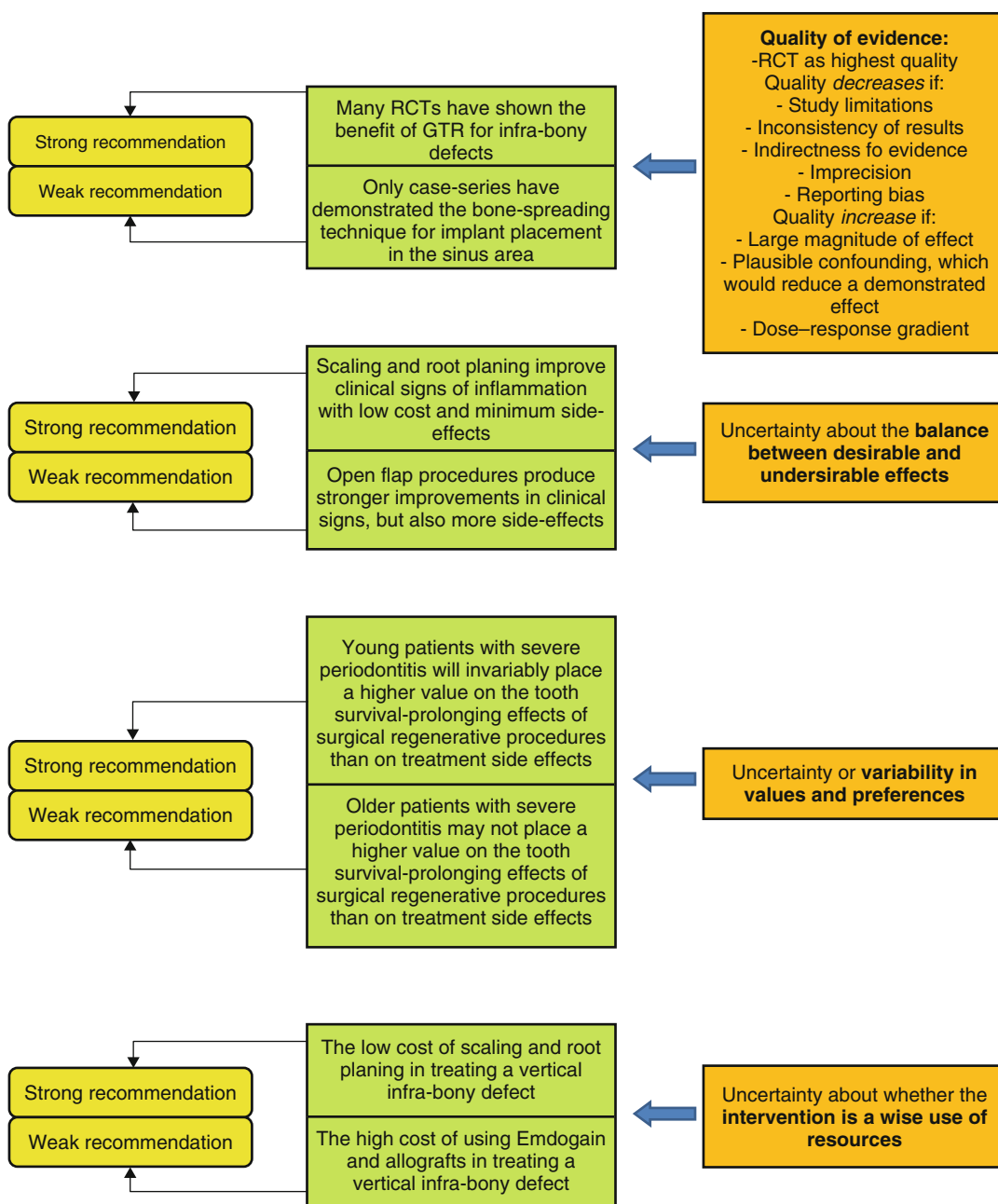


Fig. 10.2 Characteristics of the GRADE system (Faggion (2010). Reprinted with permission from Elsevier) RCT, randomized controlled trial; GTR, guided tissue regeneration

classification developed by the GRADE work group (Atkins et al. 2004). Although the GRADE system works best for treatment recommendations, it can be used to grade the quality of evidence and strength of recommendations for diagnostic tests or strategies (Schünemann et al. 2008).

10.3.3.14 Taking Account of Costs

While evidence about the effects of interventions generally comes from global research, it is necessary to take into account factors in a specific setting to inform decisions about what to do. These factors include each of the following in

Table 10.6 Characteristics of the strength of recommendation taxonomy (SORT) system (Faggion 2010) (Reprinted with permission from Elsevier)

Grade of strength of recommendation and definition	Level of quality of studies ^a	Consistency
A (based on level 1)	Level 1: SR or MA of RCTs with consistent findings; high quality individual RCTs	Consistent evidence: most studies found similar or at least coherent conclusions (coherence means that differences are explainable) or, if high-quality and up-to-date systematic reviews or meta-analyses exist, they support the recommendation
B (based on level 2)	Level 2: SR or MA of lower quality clinical trials or of studies with inconsistent findings, lower quality clinical trials, cohort study, case-control study	Inconsistent evidence: substantial variation among study findings and lack of coherence or, if high-quality and up-to-date systematic reviews or meta-analyses exist, they do not find consistent evidence in favour of the recommendation.
C (based on level 3)	Level 3: consensus guidelines, extrapolations from bench research, usual practice, opinion, disease-oriented evidence (intermediate or physiological outcomes only), or case series for studies of diagnosis, treatment, prevention, or screening	

MA meta-analysis, RCT randomized controlled trial, SR systematic review

^aQuality related to studies on treatment/prevention/screening

relationship to socioeconomic factors: the presence of effect modifiers that have been identified in the global research, baseline risk, utilization and access to care, and costs (Edejer 2006, P12).

stakeholders, and they should report the key factors that influence decisions, including those flagged in international guidelines, and the reasons for any modifications that are made (Schünemann et al. 2006e).

10.3.3.15 Applicability, Transferability, and Adaptation of Guidelines

- Resources for developing high-quality recommendations are limited. Internationally developed recommendations can facilitate access to and pooling of resources, reduce unnecessary duplication, and involve international scientists.
- Priority should be given to international health problems and problems that are important in low- and middle-income countries, where these advantages are likely to be greatest.
- Factors that influence the transferability of recommendations across different settings should be considered systematically and flagged, including modifying factors, important variation in needs, values, costs, and the availability of resources.
- Local adaptation processes should be systematic and transparent, they should involve

10.3.3.16 Structure of Reports

The Conference on Guideline Standardization (COGS) used a two-stage modified Delphi process to develop standards for reporting clinical practice guidelines (Shiffman et al. 2003). Representatives of 22 organizations active in guideline development reviewed the proposed items and commented favourably. The items were consolidated into 18 topics to create **the COGS checklist** (the main topics are (1) Overview material, (2) Focus, (3) Goal, (4) Users/setting, (5) Target population, (6) Developer, (7) Funding source/sponsor, (8) Evidence collection, (9) Recommendation grading criteria, (10) Method for synthesizing evidence, (11) Prerelease review, (12) Update plan, (13) Definitions, (14) Recommendations and rationale, (15) Potential benefits and harms, (16) Patient preferences, (17) Algorithm, (18) Implementation considerations),

which provides a framework to support comprehensive documentation of guidelines. While it is possible that some guideline developers may not include content for every item, it is suggested that they should address explicitly whether the guideline development team considered that item (Oxman et al. 2006e).

10.3.3.17 Methods of Peer Review

In the late 1990s, growing concern about variations in guideline recommendations and quality became evident. Therefore, formal methods became necessary to assess the quality of CPG. An international group of researchers from 13 countries—the Appraisal of Guidelines, Research and Evaluation (AGREE) Collaboration—developed a tool, the **AGREE Instrument**, which can be used to appraise and compare the quality of CPGs (Alonso-Coello et al. 2010). It provides a systematic framework for assessing key components of methodological quality of the guidelines, including the process of development and the quality of reporting, and intends to support both guideline developers and users. The AGREE Instrument is the only appraisal tool that has been developed and validated internationally (the AGREE Collaboration 2003). Since its publication in 2003, the AGREE Instrument has been translated into many languages and formally endorsed by several organizations (e.g. WHO Advisory Committee on Health Research) and used by many guideline development groups (Alonso-Coello et al. 2010; Edejer 2006, P12).

The AGREE Instrument contains 23 items grouped in six domains—(1) scope and purpose, (2) stakeholder involvement, (3) rigour of development, (4) clarity and presentation, (5) applicability, and (6) editorial independence—and one overall assessment item, judging whether the guideline ought to be recommended for its use in clinical practice (Table 10.7). To evaluate each item within the domains, a four-point Likert scale is used, ranging from strongly disagree to strongly agree (1–4, respectively). For the overall judgement, a three-point scale is used, ranging from not recommended to strongly recommended. At least two appraisers, but preferably four, are

needed. The score of each domain is calculated by summing the scores across the appraisers and standardizing them as a percentage of the possible maximum score (thus ranging from 0% to 100%) (Alonso-Coello et al. 2010).

As with any new assessment tool, ongoing development of the instrument is required. The AGREE Next Steps Consortium was established to conduct a program of research aimed at improving the AGREE and create the next version of the tool, the AGREE II. The consortium completed two studies (parts 1 and 2) (Brouwers et al. 2010a; Brouwers et al. 2010b). A seven-point response scale was introduced, tested its performance, and conducted a preliminary analysis of some of its measurement properties. The manual was rated by participants as appropriate, easy to use, and helpful in differentiating guidelines of varying quality, with all scores above the midpoint of the seven-point scale (Brouwers et al. 2010b). The AGREE II is available at the website of the AGREE Research Trust (www.agreetrust.org).

10.3.3.18 Planned Methods of Dissemination and Implementation

Passive dissemination of guidelines alone is not likely to adequately ensure appropriate uptake of recommendations in most circumstances. Health authorities at national or sub-national levels are better able than the WHO to tailor implementation strategies to their specific circumstances. However, they frequently lack capacity and resources to do this. The WHO headquarters and regional offices can play an important role in supporting member states in their efforts to implement recommendations by providing tools such as NorthStar, support and coordination of efforts (Fretheim et al. 2006b).

10.3.3.19 Evaluation of the Impact of the Guideline

Study designs that can be used to evaluate the impact of guidelines include randomized designs, particularly cluster randomized trials, a range of observational study designs, including interrupted time series analyses, controlled before–after stud-

Table 10.7 Domain structure for guideline quality obtained from principal components analysis, mean (SD) values of domain scores, and percentage of variance explained by each domain (item numbers represent the order in the instrument) (The AGREE Collaboration 2003) (Reprinted with permission from BMJ Publishing Group Ltd.)

	Coefficient*
Area 1: scope and purpose	
Mean percentage domain score=69.3; SD=21.3; range 16.7–97.2; % variance=4.6	
1. The overall objective of the guideline is specifically described	0.594
2. The clinical question covered by the guideline is specifically described	0.768
3. The patients to whom the guideline is meant to apply are specifically described	0.702
Area 2: stakeholder involvement	
Mean percentage domain score=36.1; SD=18.9; range 4.2–68.7; % variance=6.6	
4. The guideline development group includes individuals from all the prevalent professional groups	0.643
5. The patient's views and preferences have been sought	0.580
6. The target users of the guideline are clearly described	0.683
7. The guideline has been piloted among target users	0.471
Area 3: rigour of development	
Mean percentage domain score=40.7; SD=25.0; range 0–89.3; % variance=42.3	
8. Systematic methods were used to search for evidence	0.794
9. The criteria for selecting the evidence are clearly described	0.763
10. The methods used for formulating the recommendations are clearly described	0.750
11. The health benefits, side effects and risks have been considered in formulating the recommendations	0.689
12. There is an explicit link between the recommendations and the supporting evidence	0.753
13. The guideline has been externally reviewed by experts prior to its publication	0.589
14. A procedure for updating the guideline is provided	0.619
Area 4: clarity and presentation	
Mean percentage domain score=65.8; SD=14.1; range 37.5–91.7; % variance=8.6	
15. The recommendations are specific and unambiguous	0.716
16. The different options for management of condition are clearly presented	0.589
17. Key recommendations are easily identifiable	0.739
18. The guideline is supported with tools for application	0.640
Area 5: applicability	
Mean percentage domain score=36.9; SD=23.2; range 0–91.7; % variance=6.1	
19. The potential organisational barriers in applying the recommendations have been discussed	0.804
20. The potential cost implications of applying the recommendations have been considered	0.697
21. The guideline is supported with tools for application	0.684
Area 6: editorial independence	
Mean percentage domain score 30.3; SD=22.4; range 0–72.2	
22. The guideline is editorially independent from the funding body	
23. Conflicts of interest of guideline development members have been recorded	New item

*Coefficients from varimax rotated factor matrix.

ies and uncontrolled before–after studies. A wide variety of techniques to gather data have been used singly or in combinations, including questionnaires, interviews, observation, audit, and using routinely collected data (Oxman et al. 2006f).

Recommendations may need to be adapted to specific settings, can only be implemented

in specific settings, and their impact can only be assessed in specific settings. The WHO headquarters and regional offices, however, should support member states and those responsible for deciding and implementing policies to evaluate the impact of their policies by providing advice regarding impact assessment, practical support, and coordination of

efforts. Before–after evaluations should be used cautiously, if at all, and when there are important uncertainties regarding the effects of a policy or its implementation, randomised evaluations should be used when possible (Oxman et al. 2006f).

As presented by the Guidelines for the WHO Guidelines (http://whqlibdoc.who.int/hq/2003/EIP_GPE_EQC_2003_1.pdf), in the process of developing guidelines, the most detailed and methodologic reference is NHRMC 2000 (http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp69.pdf?q=publications/synopses/_files/cp69.pdf).

1. Refine the topics/questions

To identify the issues to be addressed, it is helpful to develop a logic and analytical frameworks' guide (Woolf et al. 1994)

2. Undertake systematic review

It is recommended that the systematic review be undertaken. (<http://hiru.mcmaster.ca/cochrane/cochrane/hbook.htm>)

After the studies have been identified and critically appraised, and the evidence synthesised, evidence should be graded.

All evidence, including that on safety, should be clearly laid out in an evidence table.

Meta-analysis should be done when the data permit.

The final results should be presented in a balance sheet

3. Draft different scenarios and develop model recommendations

To maximise representation and inputs from the different stakeholders, formal group processes should be used when possible (Murphy 1998)

4. Make recommendations for research and updating of guidelines

5. Ensure peer review

Circulate widely to experts, professional organizations, regional offices, countries

6. Make dissemination plans including plans for localization and evaluation.

7. Complete documentation of the guideline development process.

8. Submit to the steering group for approval of draft guidelines

Despite the increasing number of manuals on how to develop clinical practice guidelines

(CPGs), there remain concerns about their quality. A recent study reviewed the quality of CPGs across a wide range of health-care topics published since 1980. The authors conducted a literature search in MEDLINE to identify publications assessing the quality of CPGs with the Appraisal of Guidelines, Research and Evaluation (AGREE) Instrument. For the included guidelines in each study, the authors gathered data about the year of publication, institution, country, health-care topic, AGREE score per domain, and overall assessment. In total, 42 reviews were selected, including a total of 626 guidelines, published between 1980 and 2007, with a median of 25 CPGs. The mean scores were acceptable for the domain 'scope and purpose' (64%; 95% CI 61.9–66.4) and 'clarity and presentation' (60%; 95% CI 57.9–61.9), moderate for domain 'rigour of development' (43%; 95% CI 41.0–45.2), and low for the other domains ('stakeholder involvement' 35%; 95% CI 33.9–37.5, 'editorial independence' 30%; 95% CI 27.9–32.3, and 'applicability' 22%; 95% CI 20.4–23.9). This review showed that despite some increase in quality of CPGs over time, the quality scores as measured with the AGREE Instrument have remained moderate to low over the last two decades. This finding urges guideline developers to continue improving the quality of their products. International collaboration could help increasing the efficiency of the process (Alonso-Coello et al. 2010).

Task force researchers developed a conceptual framework—the Practice Guideline Evaluation and Adaptation Cycle—to facilitate the systematic development of a local evidence-based protocol (**Fig. 10.3**). The protocol development process involved ten major steps, as follows: (1) identify a clinical area to promote best practices; (2) establish a guideline evaluation group; (3) establish a guideline appraisal process; (4) search and retrieve existing guidelines; (5) assess the quality, currency, and content of existing guidelines; (6) adapt guidelines for local use; (7) external review; (8) finalize local guidelines; (9) official endorsement; and (10) scheduled review and revision. These steps are further explored below (Graham et al. 2005).

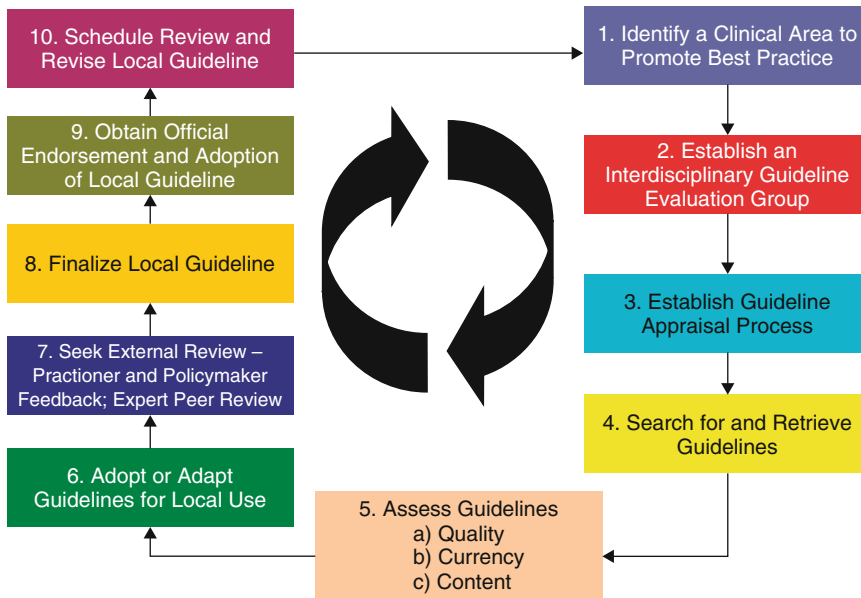


Fig. 10.3 Conceptual framework-the Practice Guideline Evaluation and Adaptation Cycle-to facilitate the systematic development of a local evidence-based protocol

(Graham et al. (2005). Reprinted with permission from Wolters Kluwer Health)

10.4 Clinical Practice Guidelines Update Process

The limitations of an arbitrary date for scheduled review include wasted resources if a full update is undertaken prematurely within a slowly evolving field. Conversely, guidelines in a rapidly evolving field may become outdated before the scheduled review (Shekelle et al. 2001b).

In brief, Shekelle et al. (2001b) identified six situations that may require a guideline to be updated (or withdrawn), including changes in (1) the available interventions, (2) the evidence on the benefits and harms of existing interventions, (3) the outcomes that are considered important, (4) the evidence that current practice is optimal, (5) the values placed on outcomes, and (6) the resources available for health care.

When the validity of 17 clinical practice guidelines published by the US Agency for Healthcare Research and Quality (AHRQ) was assessed in 2001, more than three-quarters of the AHRQ guidelines needed updating. As a general rule, it was suggested that guidelines should be reassessed for validity every 3 years (Shekelle

et al. 2001a). All American College of Physicians (ACP) clinical practice guidelines and clinical guidance statements are considered automatically withdrawn or invalid 5 years after publication or once an update has been issued (Qaseem et al. 2010).

A number of steps are required for revision (National Health and Medical Research Council Australia 1998):

- A multidisciplinary group, from disciplines similar to those of the guideline development group, should assess the guidelines to see whether there is any new evidence that should be incorporated.
- The group should also assess what has been learnt from the evaluation of the dissemination and implementation strategies and incorporate suggested improvements in the further dissemination and implementation of guidelines.
- The group should draw on current practice and experience and on national data that have been informed by the guidelines. Revision should be a coordinated activity, extending beyond the academic literature on clinical practice and health outcomes to incorporate

experience, local knowledge, and regional and national data.

- Guidelines should state the date of their development and the anticipated revision date.

Moher et al. (2007) summarized evidence of updating methods directly applicable to systematic reviews and identified five updating methods for evidence-based clinical practice guidelines (CPGs) (Table 10.8). The updating of systematic reviews and the updating of CPGs are closely related because the development of the latter often relies on scientific evidence collected and summarized in a systematic review.

10.5 Clinical Guidelines That Support Common Dental Procedures

Over the past 20 years, practice guidelines have become an increasingly popular tool for synthesis of clinical information so as to change clinical practice and improve quality of health care. Medical specialty societies have been particularly active in producing such guidelines together with agencies whose remit includes technology assessment and health-care evaluation (Grilli et al. 2000).

A systematic search for all relevant guidelines is necessary so that guidelines of which members of the group may be unaware are not overlooked. This can be done electronically, searching MEDLINE and/or EMBASE. Search terms (MeSH headings) and text words that can be used are *practice guideline*, *practice guidelines*, *clinical practice guideline*, *clinical practice guidelines*, *standards*, *consensus statement*, and *consensus* (Grilli et al. 2000). Increasingly, guideline developers are posting their guidelines on the World Wide Web, increasing the chance that they may not be indexed in commonly consulted bibliographic databases such as MEDLINE or EMBASE. For these reasons, it is now prudent to also search the World Wide Web for guidelines (Graham et al. 2002).

Clinical guidelines are considered an important tool for applying the evidence-based concept to dental practice. In 2008, Faggion analysed

whether common clinical dental procedures are supported by current freely accessible guidelines. The results of a literature search demonstrated a lack in both the number and quality of clinical dental guidelines that support common dental procedures: periodontology, oral surgery, and restorative dentistry. In periodontal field, only one guideline (published in September 2006) directly related to the dental procedure (periodontal treatment) was retrieved from the American Academy of Periodontology (AAP) site. On the same site, a position paper (published in November 2001) with characteristics of a guideline was also retrieved. However, this guideline was directed at state legislatures and agencies that regulate the practice of periodontology instead of general dental practitioners and, therefore, it was not included in the study. A further search on the AAP site retrieved two parameters of care (published in 2000 and that can be considered as guidelines) related to the treatment of chronic periodontitis. No guideline was retrieved from the National Guideline Clearinghouse (NGC), World Dental Federation (FDI), Royal College of Surgeons of England (RCSE), National Institute for Health and Clinical Excellence (NICE), and Scottish Intercollegiate Guidelines Network (SIGN) sites. On the National Library for Health site (NLH), a link to a guideline (published in November 2004, updated in October 2005) related to periodontal treatment was also retrieved. The guideline authors stated that the document was targeted at health-care professionals working within the National Health Service (UK), but did not specify if dental practitioners were included (Faggion 2008). However, a recent search of the American Academy of Periodontology (AAP) site revealed a publication from 2011 titled 'Comprehensive Periodontal Therapy: A Statement by the American Academy of Periodontology' and a 'Guidelines for the Management of Patients with Periodontal Diseases' published in 2006 (<http://www.perio.org/resources-products/posppr2.html#crpp>).

As can be seen from above, there is a lack in both number and quality of free-access clinical guidelines relating to common dental procedures. More guidelines are necessary to cover the most

Table 10.8 Methods for updating clinical practice guidelines—characteristics, strengths, and limitations (Moher et al. 2007) (Reprinted with permission from Elsevier)

Reference/organization, country	Comprehensiveness (described domains) ^a	Strengths	Limitations
Clinton (1994)/Agency for Health Care Policy and Research, ^b USA	Search strategy	(a) Minimizes publication bias by asking the public for literature	(a) Unclear what constitutes sufficient data to call a consensus meeting
	Clinical Administrative issues	(b) When sufficient data have been obtained, a meeting is held to reach consensus as to whether the guideline requires updating	(b) Unclear which individuals would be involved in the meeting
	International	(a) Efficient-targeted literature searches are used instead of a full systematic review	(a) A targeted literature search might miss important studies
Shekelle et al. (2001a, b)/AHRQ	Search strategy	(b) Assesses when an update is required based on six situations	(b) There is a stronger emphasis on when to update than how to update
	Clinical	(c) Provides a decision model to assess validity of current guidelines	(c) Including additional experts may increase costs
	Public health importance	(d) Decreases bias by using a multidisciplinary team of experts from the original guideline plus additional experts	(d) The number of outdated recommendations warranting the update of an entire guideline is arbitrary
	Economic Updating format	(e) Minimizes publication bias by asking experts for further literature	(e) Subjective judgment as to whether the guideline needs updating
	Search strategy	(a) Provides a decision process for determining the need to update	(a) Inefficient-public consultation is required for every decision
Dillon (2005)/National Institute for Clinical Excellence, UK	Clinical	(b) Decreases bias with members from the original guideline development team plus additional members to conduct the update	(b) Updates are only considered 2 or 3 years post-guideline development, which may cause out-datedness
	Economical Updating format	(c) Minimizes publication bias by obtaining data from national audits and other sources	
	Administrative issues		

Johnston et al. (2003)/Cancer Care Ontario, Canada	Search strategy Clinical Updating format Communication strategy	(a) Minimizes publication bias by obtaining data from the Cochrane library, colleagues, and hand searches	(a) Inefficient-suggests searching the literature every month, and routinely hand-searching journals
Gartlehner et al. (2004), USA	Search strategy Clinical Administrative issues	(a) Efficient-targeted literature searches are conducted instead of full systematic reviews (b) Minimizes publication bias by obtaining studies from websites of federal agencies and the Internet in general (c) The validity of the method was evaluated in an empirical study (d) Provides a schematic map of administrative and updating processes to update a guideline	(a) A targeted literature search may miss important studies (b) There is a stronger emphasis on when to update than how to update

^aDomains: search strategy, statistical method/technique, clinical (expert opinion, long-/short-term outcomes, intervention, pace of development in the field, nature of condition), public health importance (severity of condition and prevalence), economical (e.g., resource availability), updating format (preference for electronic vs. paper-based format; information steps), and administrative issues (e.g., those related to the implementation of updating process of systematic reviews)

^bNow called the Agency for Healthcare Research and Quality

common clinical dental procedures. In addition, the AGREE Instrument may help dental practitioners to objectively assess the quality of clinical guidelines (Faggion 2008).

References

- Agency for Health Care Policy and Research. Process for determining need for updates of clinical practice guidelines –AHCPR. Fed Regist. 1994;59:19723–5.
- Alonso-Coello P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, Tort S, Bonfill X, Burgers J, Schunemann H. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19:e58.
- Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, GRADE Working Group, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328:1490.
- Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401–6.
- Boyd EA, Bero LA. Improving the use of research evidence in guideline development: 4. Managing conflicts of interests. *Health Res Policy Syst*. 2006;4:16.
- Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Hanna SE, Makarski J, AGREE Next Steps Consortium. Development of the AGREE II, part 1: performance, usefulness and areas for improvement. *CMAJ*. 2010a;182:1045–52.
- Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Hanna SE, Makarski J, AGREE Next Steps Consortium. Development of the AGREE II, part 2: assessment of validity of items and tools to support application. *CMAJ*. 2010b;182:E472–8.
- Dillon A. NICE: guideline develop methods. 2005. Available at: http://www.nice.org.uk/pdf/GDM_Chapter15.pdf. Accessed 10 Dec 2005.
- Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, Ewigman B, Bowman M. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *J Am Board Fam Pract*. 2004;17(1):59–67.
- Edejer TT. Improving the use of research evidence in guideline development: incorporating considerations of cost-effectiveness, affordability and resource implications. *Health Res Policy Syst*. 2006;4:23.
- Faggion Jr CM. Clinician assessment of guidelines that support common dental procedures. *J Evid Based Dent Pract*. 2008;8:1–7.
- Faggion Jr CM. Grading the quality of evidence and the strength of recommendations in clinical dentistry: a critical review of 2 prominent approaches. *J Evid Based Dent Pract*. 2010;10:78–85.
- Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. *Health Res Policy Syst*. 2006a;4:15.
- Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 15. Disseminating and implementing guidelines. *Health Res Policy Syst*. 2006b;4:27.
- Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 5. Group processes. *Health Res Policy Syst*. 2006c;4:17.
- Garthlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP, et al. Assessing the need to update prevention guidelines: a comparison of two methods. *Int J Qual Health Care*. 2004;16:399–406.
- Graham ID, Harrison MB, Brouwers M, Davies BL, Dunn S. Facilitating the use of evidence in practice: evaluating and adapting clinical practice guidelines for local use by health care organizations. *J Obstet Gynecol Neonatal Nurs*. 2002;31:599–611.
- Graham ID, Harrison MB, Lorimer K, Piercianowski T, Friedberg E, Buchanan M, Harris C. Adapting national and international leg ulcer practice guidelines for local use: the Ontario leg ulcer community care protocol. *Adv Skin Wound Care*. 2005;18:307–18.
- Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet*. 2000;355(9198):103–6.
- Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, Raskob G, Lewis SZ, Schünemann H. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an American college of chest physicians task force. *Chest*. 2006;129:174–81.
- Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, et al. Incorporating considerations of resources use into grading recommendations. *BMJ*. 2008a;336:1170–3.
- Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is “quality of evidence” and why is it important to clinicians? *BMJ*. 2008b;336:995–8.
- Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008c;336:924–6.
- Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. Going from evidence to recommendations. *BMJ*. 2008d;336:1049–51.
- Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011a;64:383–94.
- Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of

- articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol*. 2011b;64:380–2.
- Jaeschke R, Guyatt GH, Dellinger P, Schünemann H, Levy MM, Kunz R, et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ*. 2008;337:a744.
- Johnston ME, Brouwers MC, Browman GP. Keeping cancer guidelines current: results of a comprehensive prospective literature monitoring strategy for twenty clinical practice guidelines. *Int J Technol Assess Health Care*. 2003;19:646–55.
- Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, Barrowman N. A systematic review identified few methods and strategies describing when and how to update systematic reviews. *J Clin Epidemiol*. 2007;60(11):1095–104.
- Murphy M; Black N, Lamping D. Consensus development methods and their use in clinical guideline development. *Health Technology Assessment* 1998;2: 1–88. <http://www.hta.nhsweb.nhs.uk/> (Accessed on November 2011).
- NHMRC (National Health and Medical Research Council). A guide to the development, implementation and evaluation of clinical practice guidelines. 1998. http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp30.pdf (Accessed on November 2011).
- Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 2. Priority setting. *Health Res Policy Syst*. 2006a;4:14.
- Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 7. Deciding what evidence to include. *Health Res Policy Syst*. 2006b;4:19.
- Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. *Health Res Policy Syst*. 2006c;4:20.
- Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 12. Incorporating considerations of equity. *Health Res Policy Syst*. 2006d;4:24.
- Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 14. Reporting guidelines. *Health Res Policy Syst*. 2006e;4:26.
- Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 16. Evaluation. *Health Res Policy Syst*. 2006f; 4:28.
- Qaseem A, Snow V, Owens DK, Shekelle P, Clinical Guidelines Committee of the American College of Physicians. The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. *Ann Intern Med*. 2010;153(3):194–9.
- Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Syst*. 2006a;4:21.
- Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. *Health Res Policy Syst*. 2006b;4:22.
- Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. *Health Res Policy Syst*. 2006c;4:25.
- Schünemann HJ, Fretheim A, Oxman AD, WHO Advisory Committee on Health Research. Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. *Health Res Policy Syst*. 2006d;4:13.
- Schünemann HJ, Oxman AD, Fretheim A. Improving the use of research evidence in guideline development: 6. Determining which outcomes are important. *Health Res Policy Syst*. 2006e;4:18.
- Schünemann HJ, Schünemann AH, Oxman AD, Brozek J, Glasziou P, Jaeschke R, GRADE Working Group, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008;336:1106–10.
- Shekelle PG, Eccles MP, Grimshaw JM, Woolf SH. When should clinical guidelines be updated? *BMJ*. 2001a;323:155–7.
- Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, Woolf SH. Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA*. 2001b;286:1461–7.
- Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the conference on guideline standardization. *Ann Intern Med*. 2003;139:493–8.
- Steinbrook R. Guidance for guidelines. *N Engl J Med*. 2007;356:331–3.
- The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Safe Health Care*. 2003;12:18–23.
- Woolf SH. Practice guidelines, a new reality in medicine. II. Methods of developing guidelines. *Arch Intern Med*. 1992;152(5):946–52.
- Woolf SH, DiGiuseppi CG, Atkins D, Kamerow DB. Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force. *Annu Rev Public Health*. 1996;17:511–38.
- Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. 1999;318:527–30.

Periodontal data are usually plentiful observations made in one oral cavity. In order to describe the periodontal situation, sites (gingival units) around teeth within patients or subjects are considered by using several variables. Observations may be even repeated in a longitudinal way. This is a typical hierarchical situation with lower and upper levels (Müller 2009a).

11.1 Site-Based, Tooth-Based, and Subject-Based Periodontal Evaluation

11.1.1 Clinical Perceptions

Presently, the pattern of disease progression defines the primary differentiation between the chronic and aggressive periodontitis. The current periodontal diagnostic system disregards age as a determinant in a differential diagnosis of periodontitis. The differential diagnosis between chronic and aggressive periodontitis requires examinations at more than one single time point. It is clear that the subject is the unit of observation for both conditions. In contrast, most periodontal intervention studies and clinical periodontal routines have established standard examination procedures that rely predominantly on site-based observations. Most clinical periodontal trials using the periodontal site as the unit of observation combine dichotomous (i.e. bleeding on probing, plaque scores, and tooth mobility) and numerical observations (i.e. probing depth measures and

clinical attachment loss). Such data are then often used independently (e.g. site bleeding on probing, with pocket probing depth greater than a defined threshold value) and are not merged to get information on the subject level. In the clinic, dentists make diagnostic, prognostic, and therapeutic decisions using a tooth or site as the unit of observation, automatically absorbing dichotomous and numerical information in the diagnostic process (Persson 2005).

Site-Specific Diagnostic Methods

- Traditional clinical diagnostic measures such as bleeding on probing, pocket probing depth, and clinical attachment loss have poor diagnostic and prognostic utility.
- Dental radiographs provide information on cumulative effects of periodontitis. They provide permanent records of the periodontal condition. Dental radiographs may serve as a key source for site-based periodontal diagnostics.
- Site-specific analysis of gingival crevice fluid for specific bacteria, specific markers of inflammation (enzymes and cytokines, immune markers), and specific antibody titres provides diagnostic evidence of the infection/host immune response axis. Such diagnostic methods require laboratory processing. Unlike physicians, relatively few dentists are employing laboratory-based diagnostic procedures (Persson 2005).

Subject-Based Periodontal Diagnostics

- Clinical diagnostic data (i.e. bleeding on probing pocket probing depth and radiographic data) providing subject-based information are uncommon but clearly needed.

- Subject-based information based on serum assays for antibody titres to specific pathogen in combination with microbiological data can provide differential diagnosis of periodontitis.
- Clinicians must be informed how to utilize laboratory diagnostic methods.
- Laboratory methods must attain standard diagnostic conformity and show acceptable cost-efficacy (Persson 2005).

11.1.2 Statistical Issues

The assessment of relationships between site-specific variables has been a matter of controversy because of the claim that periodontal sites within individuals can be used as independent observations in statistical models. One problem with this approach is the unreliability of the calculated type I and type II error rates. Another problem is that such inappropriate analysis may prohibit a correct assessment of causal relationships between site-specific variables. The host factor can act as an effect modifier and modulate the magnitude of the site-specific effects, and/or the host factor can act as a confounder by superimposing a patient effect on the studied site-specific effects leading to bias. As a result, site-specific biological mechanisms of disease progression may be misinterpreted (Hujoel et al. 1990).

Sites are clustered around a tooth, and teeth are clustered in individuals. It is therefore inappropriate to analyse the site- or subject-level observations using single-level, univariate statistical methods because the correlation among sites and/or teeth within an individual invalidates these statistical methods (Albandar and Goldstein 1992; Hoberman 1992; Wan et al. 2009). Consequently, statistical analysis with the assumption that the site observations are independent would generate potentially misleading results (Tu et al. 2004c).

11.2 Principles of Multilevel Modelling (MLM)

The hierarchical statistical models, or MLM, have appeared with a variety of names. In social science, they have been called **multilevel linear models**; in

biometric applications, **mixed-effects models** or **random effects models**; in economics, **random coefficient regression models**; and in statistical literature, **covariance components models** (Bryk and Raudenbush 1992; Axtelius et al. 1999).

11.2.1 The Principles of MLM

MLM has the potential to increase understanding of oral disease and health. A multilevel model makes the assumption that clustered units are broadly similar, with differences due to either random variation or discernible external influences. In MLM, the first stage of the analysis is to determine the appropriate multilevel structure. This is achieved by investigating the variation in the outcome variable attributable to each level of the proposed hierarchy. If the variance (a measure of variation) at a given level does not contribute substantially to the total variance, this level is redundant in the proposed multilevel structure and is subsequently ignored. Every possible level of a hierarchy should be considered since failure to do so can lead to a misspecified model structure, with consequent errors in the estimated model parameters (Gilthorpe et al. 2000a).

The simplest multilevel model is the *null* model, which partitions the overall variance and fits just the intercept term. Although some researchers refer to variance components models with covariates included, it is more useful to restrict the term to refer to *null* models only. Variance components models then form the basis upon which more sophisticated models are developed by subsequently incorporating covariates (Gilthorpe et al. 2000a).

11.2.2 The Advantages of Using MLM in Periodontal Research

Several valuable features of MLM can be noted. First, data can be unbalanced so that not all the observations need to be present. Second, the MLM has application to multivariate data, that is, where there are multiple outcomes recorded on the same subject. The advantage of treating multivariate structures as multilevel is that the impact

of multiple explanatory factors can be explored simultaneously. Furthermore, not all subjects require complete information for each outcome (as required for standard multivariate analyses, unless some imputing algorithm is used to estimate missing values). Third, MLM provides an opportunity to review and possibly rethink the traditional design of studies employed in much of dental research. The increased flexibility of MLM provides a more comprehensive approach to data analysis, allowing more to be asked of naturally occurring hierarchical clinical systems (Gilthorpe et al. 2000a).

11.2.3 The Disadvantages of Not Using MLM in Periodontal Research

Failure to recognize hierarchical level structures results, for example, in inflation of significance tests and confidence intervals which are too narrow. There are two principal objections to uni-level analysis when a hierarchical structure is present (Axtelius et al. 1999):

1. Firstly, aggregation of information from lower level to higher levels results in loss of information (Bryk and Raudenbush 1992; Claffey and Egelberg 1994). Any causal interpretation on a higher level must include lower levels. When information about these is discarded in the aggregation, interpretations are distorted and difficult to understand. This is the classical mistake of the ‘ecological fallacy’ (Cooley et al. 1981).
2. Secondly, disaggregation of information from higher level to lower levels results in inefficient generalizations. It involves estimating many times more coefficients than a multilevel procedure does and provides no information about the variation in the underlying population. This is known as the ‘atomistic fallacy’ (Woodhouse et al. 1995).

11.2.4 Example of a Multilevel Model

The simplest possible two-level model can be specified as two interrelated equations (Axtelius et al. 1999):

$$y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad (11.1)$$

Equation 11.1 is a micro-equation and pertains in this case to sites: The response (y_{ij}), the PD i in units j , is a function of the mean depth of PD in each of the j units (β_{0j}) plus a residual unaccounted for variation at level 1 between sites (ε_{ij}).

$$\beta_{0j} = \beta_0 + \mu_j \quad (11.2)$$

Equation 11.2 is a macro-equation and pertains to, for example, individuals: The response, β_{0j} , the mean PD in units j is a function of the overall mean depth of PD (β_0) across all units plus a residual unaccounted for variation at level 2 between individuals (μ_j).

These two equations can be combined into an overall multilevel model:

$$y_{ij} = \beta_0 + (\mu_j + \varepsilon_{ij}) \quad (11.3)$$

This so-called **null model** contains no independent variables, and the PD of a site depends only on the mean depth of all sites in all individuals (β_0) and a differential for each site (ε_j) and for each individual (μ_j). These two differentials are treated as random variables (indicated by the brackets), and their variability is summarized by two variances, σ_μ^2 and σ_ε^2 . Consequently, site PD is summarized in three parameters, the overall mean, the between-individual variance, and the within-individual, between-site variance. If individuals are not considered in PD, the ratio of the between-individual variance, σ_μ^2 , to the total variance ($\sigma_\mu^2 + \sigma_\varepsilon^2$) will be close to zero (Axtelius et al. 1999).

The null model of **Eq. 11.3** simply allows the variance to be separated into each level, but the model can readily be extended to include independent variables for both sites and individuals, as well as for an intervening level, such as tooth properties (Axtelius et al. 1999).

11.2.5 Types of Effects

A decision in specifying a multilevel model is whether the explanatory variables considered in a particular analysis have fixed or random effects:

- **Fixed effects** are those effects depicted by factors describing multiple states, where each state is distinct and of particular interest. There

is no interest in the distribution of all ‘possible’ states. Within a study, for example, the fixed effect associated with the variable subject would have as many states as there were different subjects, and each separate subject would be assessed for differences in their outcome(s) (Gilthorpe et al. 2000b).

- **Random effects** are those effects depicted by factors describing a sample of states drawn from a larger (possibly infinite) population of all ‘possible’ states. Inferences regarding the population of states are of interest. Within a study, for example, the random effect associated with the variable subject would theoretically have an infinite number of states. Each separate subject within the study would be regarded as a member of a sample drawn from the population of all ‘possible’ subjects. The impact of the population (of subjects) on an outcome would be assessed from the (random) sample of subjects (Gilthorpe et al. 2000b).

A nontechnical description of the basic principles of multilevel modelling and, as an educational example, the application of multilevel models (**the random intercept and random coefficients model**) to the data on the thickness of facial gingival were recently presented (Müller 2009a).

11.2.6 The Application of MLM to Dental Research Data

The application of the MLM method was introduced to dental research in relation to survival analysis of amalgam restorations (Gilthorpe et al. 2002), statistical methodology of meta-analysis (Lewsey et al. 2001), orthodontics (Cunningham et al. 2000; Gilthorpe and Cunningham 2000), fluorosis (Ferreira et al. 2010), dental caries (Lewsey et al. 2000; Pattussi et al. 2006; Mutsvari et al. 2010), and periodontal research (as presented below).

11.3 The MLwiN Software

MLwiN has been under development since the late 1980s, first as a command-driven DOS-based programme, MLn, and since 1998 in a fully-fledged

windows version, currently in release 1.10. It is produced by the Multilevel Models Project based within the Institute of Education, University of London, and supported largely by project funds from the UK Economic and Social Research Council. The software has been developed alongside advances in methodology and with the preparation of manuals and other training materials (Goldstein et al. 2002).

Procedures for fitting multilevel models are now available in several major software packages such as Stata, SAS, and SPSS. In addition, there are some special purpose packages, which are tailored to particular kinds of data or models (Mplus, HLM, aML, R lme4 package, Latent Gold) (Goldstein et al. 2002).

MLwiN has some particular advanced features that are not available in other packages, and it also has a user interface designed for fully interactive use (Goldstein et al. 2002).

11.4 The Application of Multilevel Modelling to Cross-Sectional Periodontal Data

Initially the data could be analysed cross-sectionally by exploring the relationship between the outcome and various risk factors at each occasion. A variance components model would confirm that significant variation existed at each level. Several risk factors would then be identified for each level, for example, presence of calculus (site level), anterior/posterior teeth (tooth level), and smoking status of the patients (subject level) (Gilthorpe et al. 2000a). This is illustrated by a series of articles by Gilthorpe et al. (2000b), Müller and Könönen (2005), Müller et al. (2005), Ciantar et al. (2005), Kim et al. (2006), Müller (2009a) and Müller (2009b), Lopez et al. (2009), Nibali et al. (2010), and Bäumer et al. (2011) where findings revealed that risk factors might ‘operate’ at more than one level (Gilthorpe et al. 2000a).

Periodontal phenotypes had been described (Seibert and Lindhe 1989) as either thick (flat, thick, and wide gingiva, interestingly associated with a square form of maxillary incisors) or thin

Table 11.1 Two-level variance components model (null model) of facial gingival thickness Significant effects in bold (modified and adapted from Müller and Könönen 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

	Estimate	95% confidence interval
<i>Fixed effects</i>		
Intercept	0.931	0.891; 0.971
<i>Random effects (variances)</i>		
Subject (σ^2_{u0})	0.008	0.001; 0.015
Tooth (σ^2_{e0})	0.183	0.166; 0.200

(scalloped, thin, and narrow gingiva, associated with slender tooth form of maxillary anterior teeth). Müller and Könönen (2005) investigated variance components of facial gingival thickness in young adults with mild gingivitis, building two-level (tooth, subject), variance components (without explanatory variables), and random intercept models with clinical site- and subject-level explanatory variables. Ordinal variables bleeding index and plaque index were encoded by design variables. **Table 11.1** presents the results of a two-level (tooth, subject) variance components model, which indicates sufficient variation across both levels. This so-called null model revealed an intercept of 0.93 (confirming mean gingival thickness) with a standard error of 0.02 mm. Subject variation amounted to 0.008 making up 4.2% of the total variance. When tooth- and subject-related clinical covariates were allowed to enter the random intercept model (**Table 11.2**), gingival thickness was positively associated with periodontal probing depth. There was also a tendency of thinner gingiva at non-bleeding sites as compared to bleeding sites on probing and sites with no or low amounts of supragingival plaque as compared to sites with plaque index of 2 or 3. At the subject level, in particular, average bleeding index and periodontal probing depth were negatively associated with gingival thickness. The random part of the model indicated that variation at both subject and tooth level was reduced. When tooth type was allowed to enter the model (**Table 11.3**), the influence of periodontal probing depth was drastically reduced but remained significant. In these young individuals with

Table 11.2 Two-level random intercept model of facial gingival thickness Significant effects in bold (Modified and adapted from Müller and Könönen 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

	Estimate	95% confidence interval
<i>Fixed effects</i>		
Intercept	0.904	0.618; 1.190
<i>Tooth level</i>		
PPD	0.333	0.280; 0.386
BI_0 ^a	−0.070	−0.158; 0.018
PI_0 ^b	−0.109	−0.177; 0.041
PI_1 ^c	−0.101	−0.182; −0.020
<i>Subject level</i>		
Average PPD	−0.154	−0.312; 0.004
Average BI	−0.343	−0.677; −0.010
Average PI	0.057	−0.052; 0.164
<i>Random effects (variances)</i>		
Subject (σ^2_{u0})	0.006	0.001; 0.011
Tooth (σ^2_{e0})	0.149	0.135; 0.163

PPD periodontal probing depth, BI bleeding index, PI plaque index

^aDesign variable encoding BI=0, reference BI±1

^bDesign variable encoding PI=0, reference PI±2

^cDesign variable encoding PI=1, reference PI±2

shallow periodontal probing depth, thickness increased, on average, by 0.067 ± 0.025 mm per millimetre probing depth. At the subject level, average bleeding index was negatively associated with gingival thickness (0.395 ± 0.149), whereas average plaque index was positively associated (0.125 ± 0.049). Tooth variation was largely reduced in this model, whereas subject variation amounted to 5.2% of the total variance (**Table 11.3**). Within its limitations, the study provided new insights into the role of subject-related factors when considering thickness of gingival tissues. In general, most of the variation of gingival thickness seems to be due to tooth type. Subject variability related to periodontal phenotype may add to the total variance, however, to a very low extent (Müller and Könönen 2005).

MLM was also used to delineate factors influencing the severity of bone loss in randomly selected orthopantomograms of adult patients seeking dental treatment. Multilevel models revealed that bone levels decreased by 0.05 mm

Table 11.3 Two-level random intercept model of facial gingival thickness adjusted for tooth type Significant effects in bold (modified and adapted from Müller and Könönen (2005) Reprinted with permission from John Wiley & Sons, Inc).

	Estimate	95% confidence interval
<i>Fixed effects</i>		
Intercept	0.673	0.409; 0.937
<i>Tooth level</i>		
PPD	0.067	0.018; 0.116
BI_0 ^a	-0.047	-0.117; 0.023
PI_0 ^b	0.010	-0.050; 0.070
PI_1 ^c	0.012	-0.055; 0.079
<i>Tooth type^d</i>		
ULI	-0.005	-0.108; 0.098
UC	-0.205	-0.309; -0.101
UP1	-0.222	-0.331; -0.113
UP2	-0.178	-0.283; -0.073
UM1	-0.150	-0.257; -0.043
UM2	-0.055	-0.165; 0.055
UM3	0.158	0.021; 0.295
LCI	-0.259	-0.363; -0.147
LLI	-0.242	-0.346; -0.138
LC	-0.259	-0.363; -0.154
LP1	-0.248	-0.393; -0.175
LP2	-0.226	-0.331; -0.121
LM1	0.084	-0.020; 0.188
LM2	0.635	0.525; 0.745
LM3	0.923	0.773; 1.073
<i>Subject level</i>		
Average PPD	0.120	-0.019; 0.259
Average BI	-0.395	-0.687; -0.103
Average PI	0.125	0.032; 0.218
<i>Random effects (variances)</i>		
Subject (σ^2_{u0})	0.005	0.001; 0.009
Tooth (σ^2_{e0})	0.091	0.083; 0.099

PPD periodontal probing depth, BI bleeding index, PI plaque index

^aDesign variable encoding BI=0, reference BI±1

^bDesign variable encoding PI=0, reference PI±2

^cDesign variable encoding PI=1, reference PI±2

^dUpper (U) and lower (L) incisors (I), canines (C), premolars (P), and molars (M); reference tooth: upper central incisor

each year of life, on average. Bone loss was more pronounced in the maxilla, especially at molars. Infrabony lesions were strongly associated with deficient restorations and peri-apical lesions (Müller et al. 2005).

11.5 The Application of Multilevel Modelling to Longitudinal Periodontal Data

Previously, burst and linear theories for periodontal disease progression were proposed based on different but limited statistical methods of analysis: the ‘linear’ (continuous-rate) model, where, overall, sites slowly and progressively lose attachment (Löe et al. 1978) and the (random) ‘burst’ model, where multiple sites show breakdown within a short period, with periods of remission that might last for months, years, or decades (Socransky et al. 1984).

Multilevel modelling provides a new approach for analysing longitudinal periodontal data, yielding a more comprehensive model. Random coefficient models were used to analyse longitudinal periodontal data consisting of repeated measures (level 1), sites (level 2), teeth (level 3), and subjects (level 4) (Gilthorpe et al. 2003). Two ‘trajectory’ models were sought to describe the values of lifetime cumulative attachment loss (LCAL) and PD respectively. A trajectory in this instance was a hypothetical ‘path’ of outcomes across all measurement occasions during the study. Any parameterization of change would suffice, provided that it facilitated the description of nonlinear transitions over time *and* provided an evaluation of how *change* (linear progression/recession) related to *acceleration/deceleration* (nonlinear progression/recession). With only three measurement occasions, trajectories were described by a quadratic expression of time through two time-dependent covariates: *linear time* (t) and *quadratic time* (t^2), where t was the duration between initial and subsequent measurements, centred on the second occasion for improved model estimation. Correlations involving the intercept differed between models, whereas correlations between both time-dependent covariates were similar (Table 11.4). These correlation structures facilitated important inferences to be made regarding longitudinal patterns of change. The large negative and highly significant correlations observed between the linear and quadratic time coefficients indicated that subjects and teeth with greater-than-average

Table 11.4 Correlations of the random coefficients for LCAL and PPD random covariate models *with* and *without* ‘risk factors’ included (modified and adapted from Gilthorpe et al. 2003) (Reprinted with permission from SAGE Publications)

	Lifetime cumulative attachment loss			Pocket probing depth		
	Intercept-t	Intercept-t ²	t – t ²	Intercept-t	Intercept-t ²	t – t ²
Random coefficient models <i>without</i> ‘risk factors’ and confounders						
Subject level	0.67*** ^a	–0.86***	–0.87***	0.26*	–0.64***	–0.72***
Tooth level	0.47***	–0.14	–0.79***	0.41***	–0.51***	–0.83***
Site level ^b	–0.05	–	–	0.19**	–	–
Random coefficient models <i>with</i> ‘risk factors’ and confounders						
Subject level	0.60**	–0.77**	–0.87***	0.04	–0.74	–0.61*
Tooth level	0.21	–0.43*	–0.43*	0.43***	–0.86***	–0.81***
Site level ^b	–0.12**	–	–	0.17**	–	–

^aStatistical significance: *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$

^bNonzero random terms at the site level were evaluated for the intercept, linear change, and the covariance between these; hence, why only one correlation *per* model could be derived at the site level

linear change experienced decelerated variation. Conversely, subjects/teeth with less-than-average linear change experienced accelerated variation. Therefore, changes in both outcomes exhibited a dynamic regression to the mean trajectory at the tooth and subject levels. Since no equilibrium was attained throughout the study, changes were thus cyclical. When considered as a multilevel system, **the ‘linear’ and ‘burst’ theories of periodontal disease progression are a manifestation of the same phenomenon: Some sites improve while others progress, in a cyclical manner** (Gilthorpe et al. 2003).

MLM was used also to investigate longitudinal relationships between the outcomes of lifetime cumulative attachment loss (LCAL) and probing depth (PD) in relation to potential risk factors for periodontal disease progression (Tu et al. 2004a, b). One hundred males (mean age 17 years) received a comprehensive periodontal examination at baseline and at 12 and 30 months. The resulting data were analysed in two stages. In stage one, the absolute levels of disease were analysed in relation to potential risk factors; in stage two, changes in disease patterns over time were analysed in relation to the same risk factors. Each approach yielded substantially different insights. The results of stage one analysis revealed that for absolute levels of disease, subject-level risk factors (covariates) had limited prediction for LCAL/PD throughout the 30-month observation period. Tooth position demonstrated a near

linear relationship for both outcomes, with disease increasing from anterior to posterior teeth. Sites with subgingival calculus and bleeding on probing demonstrated more LCAL and PD, and supragingival calculus had an apparently protective effect. Covariates had more ‘explanatory power’ for the variation in PD than for the variation in LCAL, suggesting that LCAL and PD might be generally associated with a different profile of covariates (Tu et al. 2004a). In a second stage, changes in disease (changes in lifetime cumulative attachment loss cLCAL and changes in probing depth cPD) were analysed in relation to potential risk factors or other risk markers for periodontal disease progression. Subject-level risk factors had limited predictive value for cLCAL/cPD throughout the 30-month observation period. Tooth position demonstrated a near linear relationship for both outcomes, with disease increasing from anterior to posterior teeth. Supragingival plaque had no significant effect on cLCAL/cPD, while subgingival calculus and bleeding on probing were negatively associated with cLCAL/cPD. In contrast to the outcomes LCAL/PD, supragingival calculus had no significant protective effect on cLCAL/cPD. There was no significant influence of smoking in this cohort (Tu et al. 2004b).

In a retrospective study, the value of some clinical, genetic, and radiographic variables in predicting tooth loss in periodontal patients (aged 40–60 years) treated and maintained for 10 years

Table 11.5 Odds ratios (and 95% confidence intervals) for PLI derived from a multivariate four-level model adjusted for gender (modified and adapted from Müller and Stadermann 2006) (Reprinted with permission from Springer Science + Business Media)

	Baseline	Week 8	Week 16	Week 24
Smoker	2.09 (1.79–2.44)***	1.76 (1.56–1.98)	1.71 (1.55–1.88)	1.71 (1.57–1.86)
Non-smoker	1.29 (1.10–1.51)	1.75 (1.51–2.03)	1.68 (1.50–1.87)	1.79 (1.63–1.96)

Dependent variable BOP; *** $P < 0.001$

was evaluated. Different predictor variables were then tested using a two-level statistical model (patient and tooth levels). At the patient level, these were age, gender, mean bone loss, interleukin-1 (IL-1) genotype, interaction between mean bone loss, and IL-1 genotype. At the tooth level, the variables were prosthetic restorations, molar teeth, infrabony component of the defect, bone level, and residual supporting bone. The analysis at the patient level showed that none of the considered variables proved predictive for tooth loss. At tooth level, molar teeth, the infrabony component of the defect, and residual supporting bone were significantly associated with the outcome variable and may be considered as prognostic factors for predicting tooth loss (Muzzi et al. 2006). When the prognostic value of some clinical, genetic, and radiographic variables in predicting bone level variation in these patients was assessed, it was showed that at the patient level, the prognostic factor was initial mean bone level in conjunction with a positive IL-1 genotype. At the tooth level, the prognostic factor was tooth mobility. At the site level, the significant prognostic factors were initial bone level at a site, the infrabony component of a defect, and initial probing depth at a site (Nieri et al. 2002).

Matulienė et al. (2008) investigated the influence of residual PD ≥ 5 mm and bleeding on probing (BOP) after active periodontal therapy (APT) on the progression of periodontitis and tooth loss in a retrospective cohort (172 patients) examined for 3–27 years (mean 11.3 years). Analyses were conducted using information at site, tooth, and patient levels. The association of risk factors with tooth loss and progression of periodontitis was investigated using multilevel logistic regression analysis. Residual PD ≥ 6 mm after APT represented a risk factor for both progression

of periodontitis and tooth loss during the SPT at the patient, tooth, and site level. Multiple sites (≥ 9) with residual PD ≥ 5 mm also represented a risk for further progression of periodontitis at the patient level. BOP at the site and tooth level increased the probability of tooth loss with OR 2.0 and 1.9, respectively. At the patient level, full-mouth bleeding score $\geq 30\%$ represented a risk factor for tooth loss (Matulienė et al. 2008).

Bleeding upon mechanical manipulation with, for example, a periodontal probe is regarded a rather reliable sign of gingival and periodontal inflammation induced by dental plaque (Müller and Stadermann 2006). However, several studies showed an attenuated association between the amount of supragingival plaque and gingival inflammation in smokers (Lie et al. 1998; Nair et al. 2003; Müller et al. 2002). Application of multivariate MLM for repeated measures showed that the causal relationship between bleeding on probing and amount of supragingival plaque is not affected by heavy smoking in a steady state environment.

Estimates describing this association are displayed in **Table 11.5** and were derived from a multivariate four-level model of BOP with site-specific plaque index as a covariate. With a notable exception at the outset examination when bleeding infrequently occurred, plaque consistently influenced the tendency for BOP with an odds ratio of about 1.7–1.8 for each increase in score in both smokers and non-smokers (Müller and Stadermann 2006).

Müller and Barrieshi-Nusair (2010) described a possible influence of the specific IL-1 genotype on site-specific gingival bleeding on probing (BOP) as related to strongly associated clinical variables (supragingival plaque, calculus, and PD) while considering the hierarchical

structure of the data. Fifty healthy non-smoking volunteers, 19–28 years of age, participated in this study. Clinical examinations included probing depth, BOP, plaque index (PI), and calculus. Examinations were repeated after 2 and 4 weeks, and subjects were advised not to change oral hygiene habits. Polymorphisms in the IL-1 gene cluster were assessed using a reverse hybridization assay. In multivariate multilevel models of BOP, bleeding tendency in subjects with positive IL-1 genotype was only reduced at sites with PI 0 (odds ratio of 0.98, 0.65, and 0.56 at baseline and 2 and 4 weeks, respectively; $\chi_{(3)}^2=11.946$; $P=0.008$) or 1 ($\chi_{(3)}^2=6.027$; $P=0.110$). A decreased bleeding tendency at certain tooth types in subjects who were IL-1 positive was largely related to the relative cleanliness of these areas. Random parts of the models revealed very low biserial correlations of 0.13–0.15 at the site level, whereas correlations were considerably higher (0.76–0.83) at the subject level.

11.6 The Application of Multilevel Modelling to Periodontal Treatment Response Data

It is generally believed that deeper initial pockets show more PD reduction. However, researchers have questioned whether the correlation of PD reduction and baseline PD measurement is only due to ‘**mathematical coupling** (MC)’ (Tu et al. 2002, 2005). Mathematical coupling means that part of the relationship between two variables that is due to a common component, where one of the variables is contained in the other variable or a third-dependent variable is common to both variables (Archie 1981). Archie (1981) identified four types of mathematical coupling of data. Type 1 coupling involves directional changes in two variables which are mathematically coupled. Type 2 coupling is the functional relationship between two calculated variables which have one or more common component variables. Type 3, the most common type of mathematical coupling, is direct algebraic coupling between two variables, when one or more of the variables are derived and/or calculated. Type 4 is indirect

coupling or physiological coupling. However, his classification is not always distinctive, and Tu et al. (2005) suggested that the four types of mathematical coupling should be grouped into direct (Archie’s types I, II, and III) and indirect (Archie’s type IV) mathematical coupling. Direct mathematical coupling is then defined as occurring when, in correlation or regression analysis, one variable directly or indirectly contains the whole or part of another (i.e. there is an explicit mathematical relationship between the two variables). Indirect mathematical coupling, on the other hand, refers to a relationship in correlation/regression between two variables that are functionally or physiologically associated in a biological system, without there being an explicit mathematical relationship between them (Tu et al. 2005).

Tu et al. (2004c) demonstrated the potential effect that mathematical coupling has in distorting the results from the statistical analyses of trials of dental treatment, using data from the periodontal literature on guided tissue regeneration (GTR). Three main periodontal journals were electronically and manually searched to extract the data for the clinical outcomes of pocket probing depth (PPD) and lifetime cumulative attachment loss (LCAL) in the studies using GTR. Ten out of 12 studies for PPD and 9 out of 14 for LCAL initially claimed a significant positive relationship arising from the use of simple correlation and regression techniques to analyse this association. The relationship between clinical outcomes and baseline measurements was reanalysed using Oldham’s method and the variance ratio test; after using either of the more appropriate statistical methods of adjustment, only three correlations in each group of studies remained significant. This shows that mathematical coupling caused spurious correlations between baseline disease severity and treatment effect (Tu et al. 2004c).

Regression to the mean (RTM) is a special case of MC, where coupling occurs through the ‘observed’ variables as a consequence of the measurement error. However, one can experience MC without RTM. The concepts of MC without RTM and MC that is entirely RTM are illustrated

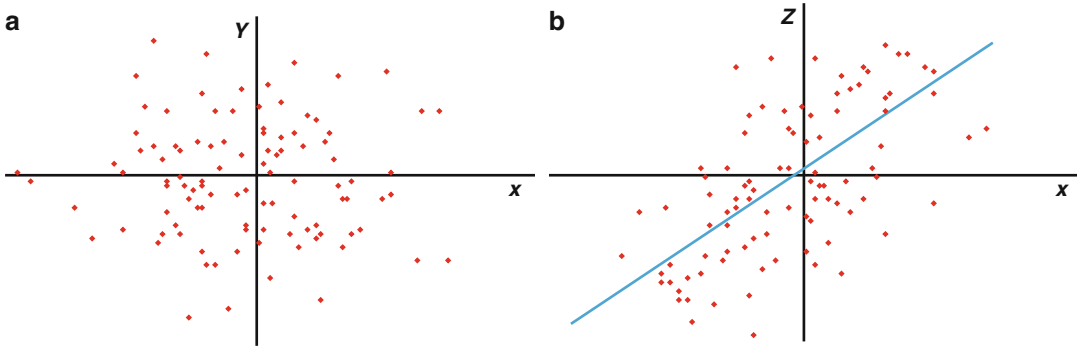


Fig. 11.1 An illustration of mathematical coupling (MC) without any effect of regression to the mean (RTM). (a) True Y values plotted against true X values, with no underlying relationship between them. (b) A spurious correlation between

true $Z (= X - Y)$ and true X is shown; the solid line is the regression of Z on X ; this represents the hypothetical situation of MC without RTM (Modified and adapted from Tu et al. (2002). Reprinted with permission from SAGE Publications)

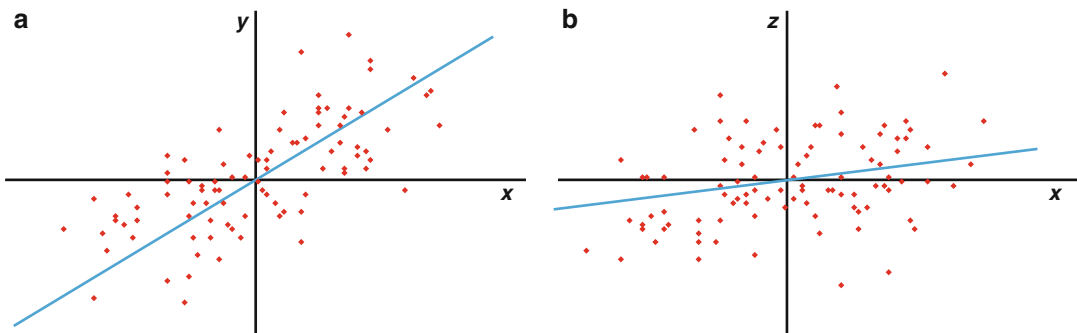


Fig. 11.2 An illustration of mathematical coupling (MC) due entirely to the effect of regression to the mean (RTM). (a) Observed y values plotted against observed x values, with the underlying relationship between them represented by the solid line. (b) A spurious correlation between

$z (= x - y)$ and x is shown; the solid line is the regression of z on x ; this represents the hypothetical situation of MC as entirely RTM (Modified and adapted from Tu et al. (2002). Reprinted with permission from SAGE Publications)

(Figs. 11.1 and 11.2) and described mathematically in Appendix (www.dentalresearch.org). Consider the situation where there is no underlying relationship between the error-free X and Y values (Fig. 11.1a). The null hypothesis for the correlation and regression coefficients between $Z (= X - Y)$ and X is that they equal $1/\sqrt{2} \approx 0.707$ and 1.0, respectively, that is, not zero (Fig. 11.1b). This shifting of the null hypothesis is a direct consequence of the coupling, and the spurious association between $Z (= X - Y)$ and X (Fig. 11.1b) is a consequence of MC without any RTM. For the situation where the slope of error-free Y values regressed on error-free X values is exactly 1, there is no underlying relationship between change and baseline value, and the correlation/regression coefficient is zero. However, in the

presence of measurement error, ‘observed’ values no longer fall on the line $Y=X$ (Fig. 11.2a), giving rise to a spurious nonzero correlation/regression coefficient between $z (= x - y)$ and x , where $x = X + e_x$ and $y = X + e_y$ (Fig. 11.2b). This is MC that is entirely RTM and is a direct consequence of the coupling of ‘observed’ values through measurement error. In reality, one does not know the relationship between X and Y , nor does one know the degree of measurement error (although the latter may be estimated from duplicate measurements made during the measurement of x and y). In practice, therefore, one is faced with a combination of ‘true’ and ‘observed’ effects. Although some studies take account of RTM, many overlook the general problem of MC (Tu et al. 2002).

In randomized controlled trials, the impact of regression to the mean on the estimation of treatment effects can be addressed by Blomqvist's formula, providing that the measurement error variance at baseline is known (or estimated); alternatively, we can simply compare the change in outcome between the treatment group and the control group. However, if the research question is whether or not the treatment works better for patients with more severe baseline disease (to be ascertained by hypothesis testing), we should apply Oldham's method to test the difference in variances between baseline and post-treatment values. In more complex scenarios, such as those with non-constant measurement errors, or the selection of subjects based on their initially high (or low) outcome values, more advanced statistical modelling is required to evaluate the correct underlying association between change and baseline, such as multilevel modelling (Tu et al. 2005; Tu and Gilthorpe 2007). In the absence of such strategies, interpretation of the results and the conclusions reached in previous studies that fail to address the MC phenomenon are questionable (Tu et al. 2002).

Multilevel analysis was applied on periodontal treatment response data in several studies (Axtelius et al. 1999; D'Aiuto et al. 2005; Hughes et al. 2006; Tomasi et al. 2007, 2008; Needleman et al. 2007; Muzzi et al. 2006; Müller 2007, 2008; Pretzl et al. 2008; Wan et al. 2009; Nibali et al. 2011).

Axtelius et al. (1999) analysed the progression of periodontal disease using analytical models that consider the level hierarchy in 22 patients with periodontitis, described as either non-responsive or responsive **to periodontal treatment**. In the null model for 'PD change' (positive values representing reduced PD after treatment) with no independent variables inserted, the subject-level variance constituted 14%, the tooth-level variance 7%, and the site-level 80% of the total sum random effect variances of the three levels. For the null model with 'final PD' (after the periodontal treatment) as the dependent variable, the corresponding values were 10%, 21%, and 70%, respectively. Similar results were reported by D'Aiuto et al. (2005) who showed that the majority of the variance in PD change after subgingival debridement was attributed to

the site level (80.4%), followed by the tooth (11.6%) and patient levels (8%).

Presence of plaque at the site level has rarely been considered as a determining factor on the outcome of **nonsurgical periodontal therapy**, but the plaque score has been used at the patient level instead. However, several studies (Axtelius et al. 1999; Tomasi et al. 2007; Wan et al. 2009) showed that the aggregated variable of plaque score on the subject level was not a significant factor, but the presence of plaque at single sites was identified as significant negative effect on the treatment outcome that was demonstrated. Also Hughes et al. (2006) and D'Aiuto et al. (2005) reported a non-significant effect of the full-mouth plaque score on the outcome of initial cause-related therapy.

Needleman et al. (2007) have shown that an intervention aimed at host modulation may effect clinical changes following nonsurgical periodontal therapy in smokers. The final improvements were no different between the test and control groups, but MLM revealed different trajectories for clinical changes. In other words, the endpoint of clinical healing was similar for both experimental groups; however, the rate of improvement was greater for the test group. Conventional statistics and MLM have very different methodologies and ask different questions of the data. The use of MLM was selected, a priori, to examine whether accounting for differences between sites within patients and at different visits might reveal differences in response to the intervention.

Using multilevel modelling analysis, Wan et al. (2009) investigated the possible factors affecting the response of nonsurgical periodontal therapy in male Chinese smokers and non-smokers in terms of both PD reduction and CAL gain. In agreement with previous reports, this study also showed that the site-level factors contributed around 70–80% of the total variance in healing outcomes, whereas tooth and subject levels only contributed the remaining 20–30%. The advantage of a multilevel approach could be identified in the difference between routine subject-level analysis, shown in **Table 11.6**, and the multilevel regression result, shown in **Table 11.7**. During the 12-month study period, the full-mouth mean PD in both groups was found to be significantly

Table 11.6 Subject-level clinical parameters over a study period (modified and adapted from Wan et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.

	Non-smokers (n = 20)				Smokers (n = 20)			
	Baseline		Months post-treatment		Baseline		Months post-treatment	
	3	6	12		3	6	12	
Full-mouth plaque %	75.45 ± 14.95	40.70 ± 17.21	32.81 ± 17.21	26.55 ± 14.19	77.36 ± 10.96	35.21 ± 23.50	26.36 ± 13.61	33.79 ± 15.07
Full-mouth BOP %	73.45 ± 21.02	42.01 ± 15.53	37.95 ± 15.40	24.92 ± 10.44	54.32 ± 13.68	32.04 ± 11.73	23.97 ± 9.65	26.91 ± 10.85
Full-mouth mean PD (mm)	2.82 ± 0.73	1.95 ± 0.42	1.82 ± 0.31	1.71 ± 0.28	2.89 ± 0.52	2.06 ± 0.37	1.99 ± 0.34	2.01 ± 0.38
Full-mouth mean CAL (mm) ^a	3.69 ± 0.97	–	–	–	3.71 ± 0.68	–	–	–
PD reduction (mm)	–	0.88 ± 0.57	1.00 ± 0.55	1.11 ± 0.69	–	0.83 ± 0.28	0.91 ± 0.28	0.89 ± 0.32
CAL gain (mm)	–	0.18 ± 0.48	0.33 ± 0.54	0.50 ± 0.52	–	0.28 ± 0.18	0.26 ± 0.21	0.31 ± 0.42
% of pocket ≥ 5.0 mm	11.43 ± 12.14	1.98 ± 2.13	1.15 ± 1.52	0.80 ± 0.94	9.98 ± 9.69	2.76 ± 3.01	2.52 ± 2.68	3.37 ± 3.24
Diseased site mean PD (mm)	5.94 ± 0.47	3.29 ± 0.57	2.89 ± 0.39	2.49 ± 0.50	5.85 ± 0.48	3.51 ± 0.71	3.46 ± 0.54	3.00 ± 0.80
Diseased site mean CAL (mm) ^a	6.86 ± 0.88	–	–	–	6.61 ± 0.64	–	–	–
Diseased site PD reduction (mm)	–	2.65 ± 0.66	3.05 ± 0.61	3.45 ± 0.62	–	2.33 ± 0.50	2.38 ± 0.57	2.84 ± 0.75
Diseased site plaque %	86.48 ± 14.67	65.24 ± 26.59	50.22 ± 26.83	42.45 ± 25.57	92.53 ± 10.24	53.23 ± 29.31	45.70 ± 23.07	58.27 ± 21.86
Diseased site BOP %	90.25 ± 15.86	65.81 ± 21.94	55.74 ± 19.73	35.86 ± 22.49	71.89 ± 19.54	44.92 ± 22.47	36.68 ± 20.88	42.42 ± 21.54

Bold font: statistically significance between groups regarding data at baseline ($P < 0.05$)
 Bold and italic fonts: statistically significance between groups after adjustment for multiple comparison ($P < 0.0017$)
BOP bleeding on probing, *CAL* clinical attachment level, *PD* probing depth
^aMeasured manually by PCP-UNC 15, Hu-Friedy Probe; other measurements of PD and CAL used Florida Probe®

Table 11.7 Final multilevel multiple regression random intercept models for reduction in PPD and gain in PAL for initially diseased sites (modified and adapted from Wan et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.

Variables	Reduction in PPD			Gain in PAL		
	3-month Estimate (SE)	6-month Estimate (SE)	12-month Estimate (SE)	3-month Estimate (SE)	6-month Estimate (SE)	12-month Estimate (SE)
Intercept	3.45 ± 0.27	3.26 ± 0.27	3.65 ± 0.28	1.71 ± 0.31	1.49 ± 0.29	2.12 ± 0.34
Subject-level						
Smoking (non-smoker versus smoker)	0.41 ± 0.20	0.79 ± 0.20	0.68 ± 0.24	-0.19 ± 0.22	-0.09 ± 0.22	-0.13 ± 0.25
Age at baseline	0.01 ± 0.02	0.02 ± 0.02	<0.01 ± 0.02	0.02 ± 0.02	0.02 ± 0.02	<0.01 ± 0.02
Number of missing teeth	<0.01 ± 0.03	<0.01 ± 0.03	0.06 ± 0.04	-0.02 ± 0.03	-0.04 ± 0.04	<0.01 ± 0.04
% of sites with plaque at baseline	<0.01 ± <0.01	<0.01 ± <0.01	-0.02 ± <0.01	-0.002 ± <0.01	-0.01 ± <0.01	-0.02 ± <0.01
% of sites with BOP at baseline	<0.01 ± <0.01	<0.01 ± <0.01	<0.01 ± <0.01	<0.01 ± <0.01	<0.01 ± <0.01	<0.01 ± <0.01
Tooth-level						
Tooth position (post. versus ant.)	-0.35 ± 0.15	-0.48 ± 0.14	-0.35 ± 0.15	-0.23 ± 0.18	-0.31 ± 0.16	-0.36 ± 0.19
Arch (lower versus upper)	-0.02 ± 0.14	-0.02 ± 0.14	0.06 ± 0.14	-0.25 ± 0.17	-0.06 ± 0.15	-0.12 ± 0.18
Site-level						
Presence of plaque at baseline	-0.55 ± 0.19	-0.44 ± 0.19	-0.45 ± 0.19	-0.48 ± 0.22	-0.45 ± 0.20	-0.25 ± 0.24
Presence of BOP at baseline	-0.21 ± 0.20	0.10 ± 0.19	-0.16 ± 0.20	-0.14 ± 0.23	-0.03 ± 0.21	-0.41 ± 0.25
Surface (lingual versus buccal)	-0.39 ± 0.13	-0.42 ± 0.13	-0.20 ± 0.13	-0.07 ± 0.15	<0.01 ± 0.14	-0.09 ± 0.16
Variance						
Subject	0.11	0.05	0.16	0.00	0.07	0.06
Tooth	0.14	0.32	0.41	0.81	0.38	0.76
Site	1.12	1.94	1.85	2.50	2.30	2.97
Total variance	1.38	2.30	2.42	3.31	2.75	3.78
% reduction in variance						
Subject	79	80	60	100	57	64
Tooth	-8	17	4	3	1	-3
Site	6	4	3	2	1	2
Total variance	9	13	11	2	5	3

Bold font: $P < 0.05$; bold and italic fonts: $P < 0.001$

^aBaseline PPD ≥ 5.0 mm

reduced when compared with the baseline. Moreover, both groups showed CAL gains compared with the baseline. However, there was no significant difference in the mean full-mouth PD reduction and the mean full-mouth CAL gain between non-smokers and smokers (**Table 11.6**). On the other hand, the multilevel regression for initially diseased sites (**Table 11.8**) showed that sites from non-smokers achieved a significantly greater PD reduction throughout the study period. The variance component models of the study clearly showed that a significant variation existed at most of the levels in the hierarchical structure at all time points. This indicates that subject-, tooth-, and site-level factors are all responsible for the outcome variations of PD reduction and change in CAL in response to nonsurgical periodontal therapy. In addition, this once more demonstrated that analysis that ignores the natural hierarchical structure of periodontal data might provide some inaccurate results.

In a recent retrospective study, Nibali et al. (2011) analysed the association between the healing of periodontal infrabony defects and baseline clinical and radiographic variables following nonsurgical treatment in a cohort of individuals treated by the same periodontist. In particular, the degree of resolution of the defect (radiographic and clinical) and its association with patient- and site-based factors were studied. Multilevel analysis was used to account for the inherent hierarchical structure of periodontal data. Following nonsurgical therapy, the horizontal (suprabony) component of the defects remained largely unchanged, while the average radiographic vertical defect depth and average defect angle changed from 3.77 mm and 37.41° at baseline to 3.08 mm and 44.11°, respectively, at re-evaluation ($P < 0.001$ for both parameters). The results from the multilevel analysis (**Table 11.8**) showed that greater radiographic defect depth reduction was found in sites with greater initial defect depth (0.28 mm, 95% CI = 0.08–0.48 mm, $P = 0.005$) and in patients with the use of adjunctive antibiotics (systemic $n = 11$, local $n = 1$) (1.10 mm, 95% CI = 0.29–1.90 mm, $P = 0.007$). Current smokers had less radiographic defect depth reduction (–0.72 mm, 95% CI = –1.39 to 0.05 mm, $P = 0.034$) than

non-smokers. Defect angle changes (defined as re-evaluation minus baseline) were negatively associated with the initial defect angle (–0.361, 95% CI = –0.71 to –0.011, $P = 0.045$), that is, lesions with greater angles on average showed smaller increases in defect angles, and lesions adjacent to premolars (–8.98°, 95% CI = –15.86° to –2.10°, $P = 0.011$) showed less increase in the defect angle than those adjacent to anterior teeth. Patients treated with adjunctive antibiotics showed greater increases in defect angles (10.25°, 95% CI = 1.79–18.70°, $P = 0.018$).

Müller (2008) modelled the alterations of mucosal thickness after the implantation of a bio-absorbable membrane for surgical root coverage employing guided tissue regeneration. Periodontal conditions around 31 recession sites in 14 patients were assessed for up to 12 months after surgery. Mucosal thickness was modelled in a multivariate, three-level (occasion, tooth, subject) time series model. The amount of root coverage was studied in a bivariate multilevel model of change and mean recession to avoid mathematical coupling. Predictions of gingival thickness were, at the outset, strongly related to baseline gingival width. At maxillary teeth, gingival thickness at all measurement locations peaked 3 months after surgery with negative relations to baseline gingival width. Thereafter, thickness gradually decreased but remained higher (0.3–0.5 mm, 95% confidence interval 0.05–0.9 mm) than before surgery, while positive correlations with baseline gingival width were re-established. At mandibular teeth, gingival thickness did not change so dramatically, while thickness of lining mucosa underwent similar changes as at maxillary teeth (Müller 2008).

Multivariate multilevel models used to study longitudinal associations between plaque and gingival bleeding indicated that (I) 6 weeks after introducing 0.3% triclosan/2% copolymer-containing toothpaste, the correlation between mean plaque index and bleeding scores was apparently attenuated in young female non-smokers, while the introduction of fluoride-containing toothpaste without triclosan did not alter this relationship; (II) whereas the likelihood of BOP at the site level was decreased by the test toothpaste, plaque index and presence of supragingival calculus were not

Table 11.8 Results from multilevel linear regression for association with reduction in radiographic defect depth and defect angle (modified and adapted from Nibali et al. (2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Outcome	Defect depth change			Defect angle change		
	Coefficients	95% confidence interval	P-value	Coefficients	95% confidence interval	P-value
Fixed effects						
Subject factors						
Smoking (current)	-0.72	(-1.39, -0.05)	0.034	-4.47	(-11.62, 2.69)	0.221
Gender (male)	0.16	(-0.57, 0.89)	0.673	2.19	(-5.36, 9.74)	0.570
Age	0.002	(-0.03, 0.03)	0.886	0.16	(-0.18, 0.50)	0.357
AgP diagnosis (versus CP)	0.03	(-1.17, 1.23)	0.962	-1.65	(-14.44, 11.14)	0.800
Site factors						
Previous endodontic treatment	-0.24	(-1.25, 0.77)	0.64	-0.35	(-10.11, 9.41)	0.944
Pre-molar (versus anterior)	-0.73	(-1.47, 0.14)	0.055	-8.98	(-15.86, -2.10)	0.011
Molar (versus anterior)	-0.47	(-1.25, 0.30)	0.229	-0.06	(-7.21, 7.09)	0.986
Initial defect depth	0.28	(0.08, 0.48)	0.005	-0.53	(-2.38, 1.32)	0.575
Initial defect angle	0.01	(-0.02, 0.05)	0.501	-0.36	(-0.71, -0.01)	0.045
Visible calculus (on radiograph)	0.44	(-0.55, 1.44)	0.382	7.80	(-1.71, 17.31)	0.108
Treatment factors						
Multiple sessions (versus 1-day therapy)	0.48	(-0.25, 1.21)	0.200	5.34	(-2.40, 13.09)	0.176
Adjunctive antibiotics	1.10	(0.29, 1.90)	0.007	10.25	(1.79, 18.70)	0.018
Random effects						
Level-2 variance	0.39	(0.66-2.36)		7.49	(4.0-14.0)	
Level-1 variance	1.44	(1.21-1.71)		12.64	(10.42-14.34)	

Only the final model is reported (other subject and site factors were tested previously and discarded from the model as non-significant). The regression coefficients from multilevel linear regression analysis are reported in the first column, followed by confidence intervals and p values for each investigated factor

CP chronic periodontitis, AgP aggressive periodontitis

affected; (III) a site-specific attenuation of the causal association between plaque index scores and BOP in subjects using triclosan in toothpaste could already be discerned after 6 weeks; and (IV) not at least, multilevel modelling of the abundance of observations usually acquired in periodontal examinations in subjects has definitely the potential to provide deeper insights into pathogenetic and therapeutic mechanisms at the level of interest, the periodontal site, while concomitant consideration of random parameters, that is, variances and covariances, at the site, tooth, and subject levels is allowed as well (Müller et al. 2006).

The main advantage of multilevel modelling in periodontal research is the possibility of unbiased dealing with the true, hierarchical, structure of the data. New study designs should take into consideration the tremendous power of these techniques (Müller 2009a).

References

- Albandar JM, Goldstein H. Multi-level statistical models in studies of periodontal diseases. *J Periodontol.* 1992;63:690–5.
- Archie Jr JP. Mathematic coupling of data: a common source of error. *Ann Surg.* 1981;193:296–303.
- Axtelius B, Söderfeldt B, Attström R. A multilevel analysis of factors affecting pocket probing depth in patients responding differently to periodontal treatment. *J Clin Periodontol.* 1999;26:67–76.
- Bäumer A, Pretzl B, Cosgarea R, Kim TS, Reitmeir P, Eickholz P, Dannewitz B. Tooth loss in aggressive periodontitis after active periodontal therapy: patient-related and tooth-related prognostic factors. *J Clin Periodontol.* 2011;38:644–51. doi:10.1111/j.1600-051X.2011.01733.x.
- Bryk AS, Raudenbush SW. Hierarchical linear models. Applications and data analysis methods. Newbury Park: Sage Publications; 1992.
- Ciantar M, Gilthorpe MS, Hurel SJ, Newman HN, Wilson M, Spratt DA. *Capnocytophaga* spp. in periodontitis patients manifesting diabetes mellitus. *J Periodontol.* 2005;76:194–203.
- Claffey N, Egelberg J. Clinical characteristics of periodontal sites with probing attachment loss following initial periodontal treatment. *J Clin Periodontol.* 1994;21:670–9.
- Cooley W, Bond L, Mao BJ. Analysing multilevel data. In: Berk R, editor. Educational evaluation methodology. Baltimore: Johns Hopkins University Press; 1981.
- Cunningham SJ, Gilthorpe MS, Hunt NP. Are orthognathic patients different? *Eur J Orthod.* 2000;22:195–202.
- D’Aiuto F, Ready D, Parkar M, Tonetti MS. Relative contribution of patient-, tooth-, and site-associated variability on the clinical outcomes of subgingival debridement. I. Probing depths. *J Periodontol.* 2005;76:398–405.
- Ferreira EF, Vargas AM, Castilho LS, Velásquez LN, Fantinel LM, Abreu MH. Factors associated to endemic dental fluorosis in Brazilian rural communities. *Int J Environ Res Public Health.* 2010;7:3115–28.
- Gilthorpe MS, Cunningham SJ. The application of multilevel, multivariate modelling to orthodontic research data. *Community Dent Health.* 2000;17:236–42.
- Gilthorpe MS, Griffiths GS, Maddick IH, Zamzuri AT. The application of multilevel modelling to periodontal research data. *Community Dent Health.* 2000a;17:227–35.
- Gilthorpe MS, Mayhew MT, Bulman JS. Multilevel survival analysis of amalgam restorations amongst RAF personnel. *Community Dent Health.* 2002;19:3–11.
- Gilthorpe MS, Maddick IH, Petrie A. Introduction to multilevel modelling in dental research. *Community Dent Health.* 2000b;17:222–6.
- Gilthorpe MS, Zamzuri AT, Griffiths GS, Maddick IH, Eaton KA, Johnson NW. Unification of the “burst” and “linear” theories of periodontal disease progression: a multilevel manifestation of the same phenomenon. *J Dent Res.* 2003;82:200–5.
- Goldstein H, Browne W, Rasbash J. Multilevel modelling of medical data. *Stat Med.* 2002;21:3291–315.
- Hoberman D. Periodontal sites or patients as the experimental unit. *J Periodontal Res.* 1992;27:417–9.
- Hughes FJ, Syed M, Koshy B, Marinho V, Bostanci N, McKay IJ, Curtis MA, Croucher RE, Marcenes W. Prognostic factors in the treatment of generalized aggressive periodontitis: I. Clinical features and initial outcome. *J Clin Periodontol.* 2006;33:663–70.
- Hujoel PP, Loesche WJ, DeRouen TA. Assessment of relationships between site-specific variables. *J Periodontol.* 1990;61:368–72.
- Kim CK, Choi SH, Kim TS, Kaltschmitt J, Eickholz P. The infrabony defect and its determinants. *J Periodontal Res.* 2006;41:498–502.
- Lewsey JD, Gilthorpe MS, Bulman JS, Bedi R. Is modelling dental caries a “normal” thing to do? *Community Dent Health.* 2000;17:212–7.
- Lewsey JD, Gilthorpe MS, Gulabivala K. An introduction to meta-analysis within the framework of multilevel modelling using the probability of success of root canal treatment as an illustration. *Community Dent Health.* 2001;18:131–7.
- Lie MA, Timmerman MF, van der Velden U, van der Weijden GA. Evaluation of 2 methods to assess gingival bleeding in smokers and non-smokers in natural and experimental gingivitis. *J Clin Periodontol.* 1998;25:695–700.
- Lopez R, Frydenberg M, Baelum V. Contextual effects in the occurrence of periodontal attachment loss and necrotizing gingival lesions among adolescents. *Eur J Oral Sci.* 2009;117:547–54.
- Löe H, Anerud A, Boysen H, Smith M. The natural history of periodontal disease in man. The rate of periodontal disease destruction before 40 years of age. *J Periodontol* 1978;49:607–620.
- Matuliene G, Pjetursson BE, Salvi GE, Schmidlin K, Brägger U, Zwahlen M, Lang NP. Influence of residual

- pockets on progression of periodontitis and tooth loss: results after 11 years of maintenance. *J Clin Periodontol*. 2008;35:685–95.
- Müller HP. A bivariate multi-level model, which avoids mathematical coupling in the study of change and initial periodontal attachment level after therapy. *Clin Oral Investig*. 2007;11:307–10.
- Müller HP. Modeling mucosal dimensions after implantation of a bio-absorbable membrane for surgical root coverage. *Clin Oral Investig*. 2008;12:249–55.
- Müller HP. Dealing with hierarchical data in periodontal research. *Clin Oral Investig*. 2009a;13:273–8.
- Müller HP. Multilevel modeling of gingival bleeding on probing in young adult carriers of non-JP2-like strains of *Aggregatibacter actinomycetemcomitans*. *Clin Oral Investig*. 2009b;13:171–8.
- Müller HP, Barrieshi-Nusair KM. Site-specific gingival bleeding on probing in a steady-state plaque environment: influence of polymorphisms in the interleukin-1 gene cluster. *J Periodontol*. 2010;81:52–61.
- Müller HP, Könönen E. Variance components of gingival thickness. *J Periodontol Res*. 2005;40:239–44.
- Müller HP, Stadermann S. Multivariate multilevel models for repeated measures in the study of smoking effects on the association between plaque and gingival bleeding. *Clin Oral Investig*. 2006;10:311–6.
- Müller HP, Stadermann S, Heinecke A. Longitudinal association between plaque and gingival bleeding in non-smokers and smokers. *J Clin Periodontol*. 2002;29:287–94.
- Müller HP, Ulbrich M, Heinecke A. Alveolar bone loss in adults as assessed on panoramic radiographs. (II) multilevel models. *Clin Oral Investig*. 2005;9:105–10.
- Müller HP, Barrieshi-Nusair KM, Könönen E, Yang M. Effect of triclosan/copolymer-containing toothpaste on the association between plaque and gingival bleeding: a randomized controlled clinical trial. *J Clin Periodontol*. 2006;33:811–8.
- Mutsvari T, Lesaffre E, García-Zattera MJ, Diya L, Declerck D. Factors that influence data quality in caries experience detection: a multilevel modelling approach. *Caries Res*. 2010;44:438–44.
- Muzzi L, Nieri M, Cattabriga M, Rotundo R, Cairo F, Pini Prato GP. The potential prognostic value of some periodontal factors for tooth loss: a retrospective multilevel analysis on periodontal patients treated and maintained over 10 years. *J Periodontol*. 2006;77:2084–9.
- Nair P, Sutherland G, Palmer RM, Wilson RF, Scott DA. Gingival bleeding on probing increases after quitting smoking. *J Clin Periodontol*. 2003;30:435–7.
- Needleman I, Suvan J, Gilthorpe MS, Tucker R, St George G, Giannobile W, Tonetti M, Jarvis M. A randomized-controlled trial of low-dose doxycycline for periodontitis in smokers. *J Clin Periodontol*. 2007;34(4):325–33.
- Nibali L, Donos N, Farrell S, Ready D, Pratten J, Tu YK, D'Aiuto F. Association between interleukin-6-174 polymorphism and *Aggregatibacter actinomycetemcomitans* in chronic periodontitis. *J Periodontol*. 2010;81:1814–9.
- Nibali L, Pometti D, Tu Y-K, Donos N. Clinical and radiographic outcomes following non-surgical therapy of periodontal infrabony defects: a retrospective study. *J Clin Periodontol*. 2011;38:50–7.
- Nieri M, Muzzi L, Cattabriga M, Rotundo R, Cairo F, Pini Prato GP. The prognostic value of several periodontal factors measured as radiographic bone level variation: a 10-year retrospective multilevel analysis of treated and maintained periodontal patients. *J Periodontol*. 2002;73:1485–93.
- Pattussi MP, Hardy R, Sheiham A. The potential impact of neighbourhood empowerment on dental caries among adolescents. *Community Dent Oral Epidemiol*. 2006;34:344–50.
- Persson GR. Site-based versus subject-based periodontal diagnosis. *Periodontol* 2000. 2005;39:145–63.
- Pretzl B, Kaltschmitt J, Kim TS, Reitmeir P, Eickholz P. Tooth loss after active periodontal therapy. 2: tooth-related factors. *J Clin Periodontol*. 2008;35:175–82.
- Seibert J, Lindhe J. Esthetics and periodontal therapy. In: Lindhe J, editor. *Textbook of clinical periodontology*. 2nd ed. Copenhagen: Munksgaard; 1989. p. 477–514.
- Socransky SS, Haffajee AD, Goodson JM, Lindhe J. New concepts of destructive periodontal disease. *J Clin Periodontol* 1984;11:21–32.
- Tomasi C, Leyland AH, Wennström JL. Factors influencing the outcome of non-surgical periodontal treatment: a multilevel approach. *J Clin Periodontol*. 2007;34:682–90.
- Tomasi C, Koutouzis T, Wennström JL. Locally delivered doxycycline as an adjunct to mechanical debridement at retreatment of periodontal pockets. *J Periodontol*. 2008;79:431–9.
- Tu YK, Gilthorpe MS. Revisiting the relation between change and initial value: a review and evaluation. *Stat Med*. 2007;26:443–57.
- Tu YK, Gilthorpe MS, Griffiths GS. Is reduction of pocket probing depth correlated with the baseline value or is it “mathematical coupling”? *J Dent Res*. 2002;81:722–6.
- Tu YK, Gilthorpe MS, Griffiths GS, Maddick IH, Eaton KA, Johnson NW. The application of multilevel modeling in the analysis of longitudinal periodontal data – part I: absolute levels of disease. *J Periodontol*. 2004a;75:127–36.
- Tu YK, Gilthorpe MS, Griffiths GS, Maddick IH, Eaton KA, Johnson NW. The application of multilevel modeling in the analysis of longitudinal periodontal data – part II: changes in disease levels over time. *J Periodontol*. 2004b;75:137–45.
- Tu YK, Maddick IH, Griffiths GS, Gilthorpe MS. Mathematical coupling can undermine the statistical assessment of clinical research: illustration from the treatment of guided tissue regeneration. *J Dent*. 2004c;32:133–42.
- Tu YK, Baelum V, Gilthorpe MS. The problem of analysing the relationship between change and initial value in oral health research. *Eur J Oral Sci*. 2005;113:271–8.
- Wan CP, Leung WK, Wong MC, Wong RM, Wan P, Lo EC, Corbet EF. Effects of smoking on healing response to non-surgical periodontal therapy: a multilevel modelling analysis. *J Clin Periodontol*. 2009;36:229–39.
- Woodhouse G, Rasbash J, Goldstein H, Yang M. Introduction to multilevel modelling. In: Woodhouse G, editor. *A guide to MLn for new users*. London: Institution of Education, University of London; 1995. p. 9–58.

Periodontitis is an inflammatory disease of bacterial origin that results in the progressive destruction of the tissues that support the teeth, specifically the gingiva, periodontal ligament, and alveolar bone. Although there have been significant advances in the understanding of the cause and pathogenesis of periodontal disease over the past 40 years, the traditional methods by which clinicians diagnose periodontal disease have remained virtually unchanged. The diagnosis of periodontal disease relies almost exclusively on clinical parameters and traditional dental radiography. These traditional diagnostic tools have some significant shortcomings. Demonstrating progressive loss of periodontal support requires longitudinal assessment. Current diagnostic methodologies do not enable us to accurately predict which periodontal sites, teeth, or individuals are susceptible to further periodontal breakdown. Given the limitations of current diagnostic tools, researchers are working to develop techniques that address some of these inadequacies. New approaches and technologies that are being developed to improve the diagnosis of periodontal disease are here presented (Wolf and Lamster 2011; Sanz et al. 2011; Agrawal et al. 2012).

12.1 Clinical Methods for Periodontal Diagnosis

12.1.1 Subgingival Temperature

Attempts have been made to develop other measures of periodontal inflammation and disease

activity that may be useful when the usual clinical signs are unreliable. One such measure is subgingival temperature. This may be recorded using a commercially available device known as the PeriTemp™ (Abiodent, Danvers, MA). This may also be used to measure sublingual temperature and thereby enable calculation of the temperature differential (ΔT) between this and subgingival temperature. The temperature differential is useful because it allows differences in core temperature between individuals to be taken into account. Studies on subgingival temperature have demonstrated that the subgingival temperature at diseased sites are raised compared with healthy sites and that a natural anterior/posterior temperature gradient exists (Kung et al. 1990; Fedi and Killoy 1992). Haffajee et al. (1992a) confirmed the relationship between raised subgingival temperature and disease, and the anatomical temperature gradient within dental arches (posterior sites warmer than anterior sites). In addition, mandibular sites were reported to be warmer than maxillary sites (Maiden et al. 1998).

In subjects with periodontal disease, higher subgingival temperatures were associated with sites which harboured plaque, were red, bled on probing, exhibited suppuration, have deeper pockets, and have more attachment loss (Haffajee et al. 1992a; Holthuis et al. 1981). The elevated subgingival temperature appears to be also a good indicator of the risk of subsequent attachment loss at shallow sites. The 8.9% of sites with shallow pockets which bled on probing and had a clinical attachment level loss >2 mm (detected in

the following 2 months) had higher subgingival temperatures compared with the shallow pockets with the lower risk (1.4%) non-bleeding sites which had lower subgingival temperatures (Haffajee et al. 1992b). Subgingival temperature is correlated with the GCF enzymes, neutrophil elastase, myeloperoxidase, and beta-glucuronidase which have previously been identified to be associated with sites of inflammation (Wolff et al. 1997).

A study of subjects with moderate periodontitis (Haffajee et al. 1992c) reported associations between certain subgingival species (*B. intermedius*, *P. micros*, *P. gingivalis*, and *A. actinomycetemcomitans*) and site and subject subgingival temperatures. Wolff et al. (1997) also found statistically significant correlations between subgingival temperature and two of the periodontitis-associated plaque microorganisms, *E. corrodens* and *F. nucleatum*.

Differences in the temperature differentials between subgingival temperature and sublingual temperature (ΔT) were also reported in smokers compared to nonsmokers (Dinsdale et al. 1997). For healthy sites, smokers had a mean ΔT 0.2°C lower than nonsmokers (warmer sites), and for diseased sites, ΔT was 0.3°C higher in smokers (cooler sites). Similarly, it was shown that for maxillary buccal anterior sites, there is a decrease of temperature differentials with an increase of probing depth at bleeding sites for both smokers and nonsmokers. Smokers were found to have significantly increased ΔT (suggesting lower subgingival temperatures) compared to nonsmokers, at probing depths of 2, 3, 4, and 5 mm (Trikilis et al. 1999).

The repeatability of subgingival temperature measurements was previously determined by Haffajee et al. (1992a), who found the accuracy and reproducibility of the Perio Temp System to be $\pm 0.1^\circ\text{C}$. Further work by Niederman and Kent (1993) found that the Perio Temp System can accurately record temperature differences between 33°C and 37°C with a standard deviation of less than 0.09°C. The repeatability of subgingival temperature was found to be acceptable during an examiner calibration phase by Trikilis et al. (1999).

12.1.2 Bleeding on Probing

Clinical evaluation of the degree of gingival inflammation includes assessment of the redness and swelling of the gingiva along with assessment of gingival bleeding (Figs. 12.1 and 12.2). Although the earliest clinical signs of gingivitis consist of colour and texture changes, there may be underlying structural alterations without corresponding clinical signs. Gingival bleeding is related to the persistent presence of plaque on the teeth and is regarded as a sign of the associated inflammatory response (Sanz et al. 2009). Studies previously referred indicate that gingiva that does not bleed upon probing demonstrates a more dense collagenous connective tissue and a smaller inflammatory infiltrate than tissue that do bleed (Greenstein 1984).

Primarily for use in clinical studies, several indices or grading systems have been developed to formally characterize the appearance of bleeding on gentle manipulation of inflamed gingiva. Since the amount of pressure used during probe insertion can influence whether or not a site bleeds, it may be advisable to use controlled-force pressure-sensitive periodontal probes when bleeding indices are taken (Armitage et al. 1994; Müller and Heinecke 2002).

In both cross-sectional and longitudinal studies of young adults with plaque-induced gingivitis, it has been observed that bleeding upon probing is only weakly associated with supragingival plaque (Müller et al. 2000, 2002). It has been speculated that gingival bleeding may be influenced by several independent factors other than plaque. Even light smoking appears to be an independent risk factor for gingivitis, whereas in multivariable models adjusting for smoking, tooth type, and clinical variables, gingival dimensions did not influence bleeding tendency (Müller and Heinecke 2002).

Although bleeding on probing may have a limited predictive value for disease progression (Haffajee et al. 1983; Badersten et al. 1985; Armitage et al. 1994), its absence indicates periodontal stability with high probability (Lang et al. 1990). However, this fact may not be true in heavy smokers (Dietrich et al. 2004; Warnakulasuriya et al. 2010). Cigarette

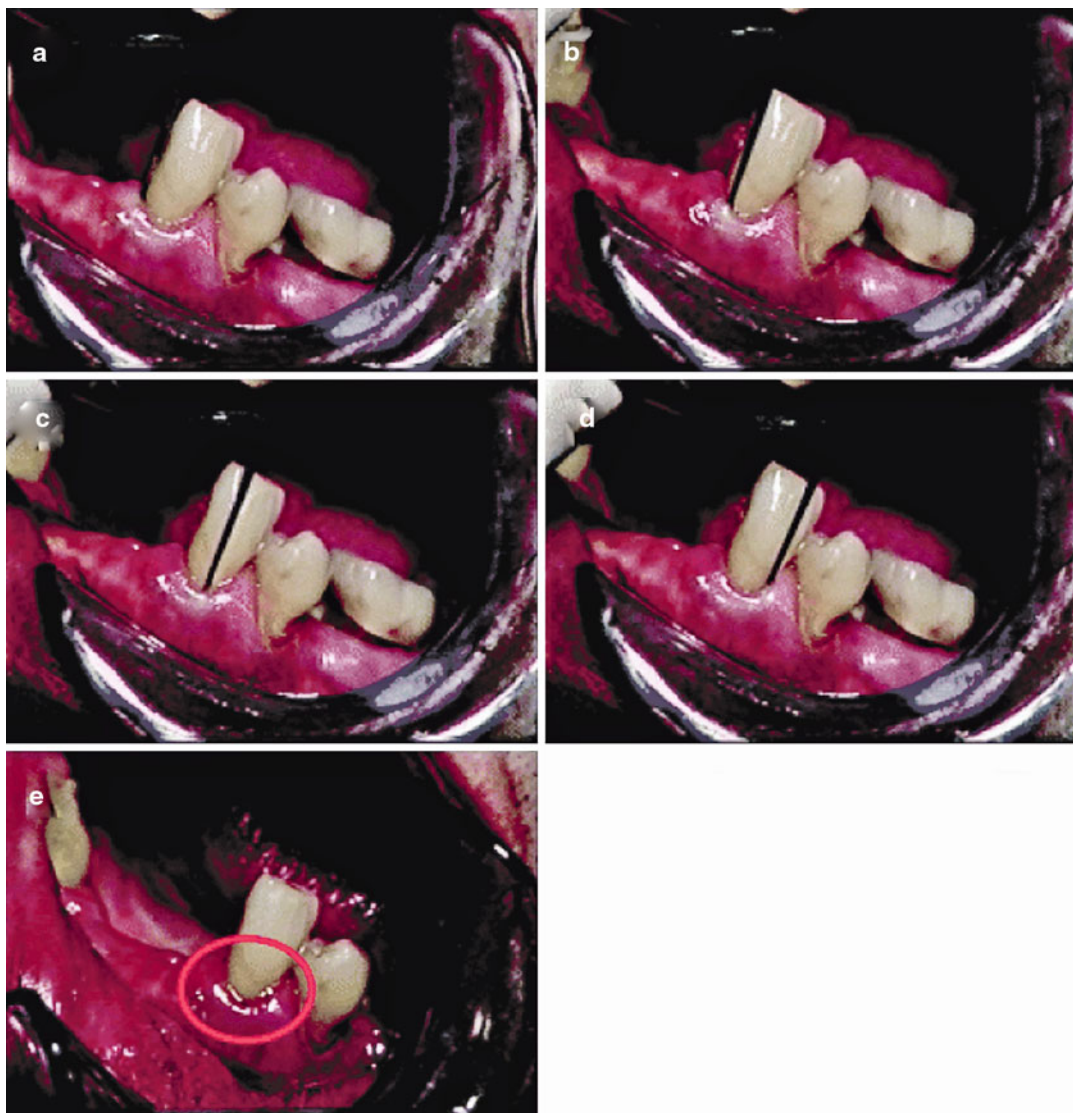


Fig. 12.1 Assessment of marginal gingivitis by a ‘gingival sweep’ from proximal surface to proximal surface. No/minimal bleeding was noted following the gingival

sweep (*red circle*) (Scott and Singer 2004. Reprinted with permission from John Wiley & Sons, Inc.)

smoking is not only a risk factor for the development of periodontitis that also reduces the response to conventional treatments of periodontitis, but—by masking the inflammatory response—smoking also makes clinical decision-making about the diagnosis and treatment planning for this disease much more difficult. It is also apparent that while smoking may suppress gingival angiogenesis, the mechanisms by which cigarette smoking dampens periodontal

inflammation are not yet fully understood (Scott and Singer 2004).

The finding that the inhibition of gingival inflammation, as measured by bleeding on probing, is reversible (Nair et al. 2003b; Morozumi et al. 2004) has important implications in clinical practice. Oral health professionals should advise patients who are planning to quit of the possibility of inflammatory recovery in the periodontal tissues and, therefore, increased gingival bleeding.

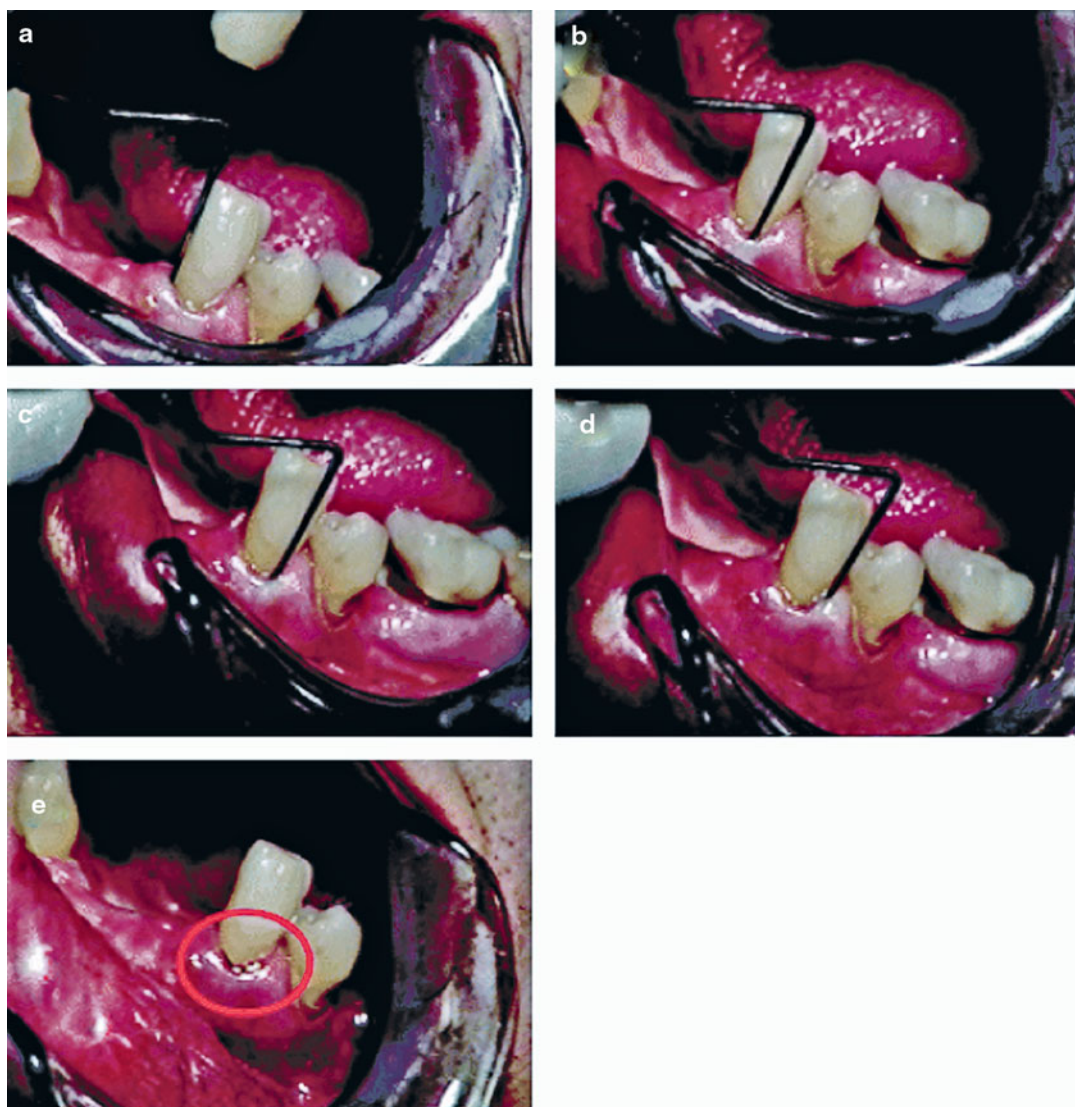


Fig. 12.2 Assessment of bleeding on probing to pocket depth. Probing was performed at the same site, and in the same patient, as shown in Fig. 12.1 (red circle) (Scott and Singer 2004. Reprinted with permission from John Wiley & Sons, Inc.)

This would prevent anxiety over these gingival changes, which could compromise cessation attempts by patients (Scott and Singer 2004).

12.1.3 Assessment of Loss of Periodontal Attachment

The word *probe* is derived from the Latin word *Probo*, which means ‘to test’. Periodontal probes are used primarily to detect and measure periodontal pockets and clinical attachment loss. In

addition, they are used to locate calculus; measure gingival recession, width of attached gingiva, and size of intraoral lesions; identify tooth and soft tissue anomalies; locate and measure furcation involvements; and determine mucogingival relationships and bleeding tendencies (Ramachandra et al. 2011).

12.1.3.1 Periodontal Terminology

The body of knowledge that was the basis of the science and therapy of periodontics can be found in the literature of the 1950s and 1960s. Even

with workshops designed in part to arrive at consensus to preclude confusion, the literature is replete with misleading use of nomenclature. The lack of attention to the appropriate use of terms is, at best, confusing to the student, researcher, and clinician. In the worst case, it may result in poor performance on examinations and improper evaluation of techniques and materials. Many periodontal terms are used inappropriately or are not completely understood probably because these words have not been adequately described in the literature (Weinberg and Eskow 2003).

It was recommended to use the term **probing depth** instead of probing pocket depth or pocket depth because it is not known if disease or health is present. For example, it is incorrect to say the ‘pocket depth’ on the mesial surface of a tooth is 2 mm if there is no sign of disease and therefore a pocket does not exist. **Clinical attachment level (CAL)** should not be confused with the term connective tissue attachment loss. Clinical attachment level is the distance from the CEJ (a stationary, stable point of reference) to the most apical extent of the periodontal probe penetration. It is the clinical approximation of the level of soft tissue attachment (junctional epithelium) to the tooth surface. Clinical attachment level simply gives information on the level of the tissue attachment to the tooth (and it is not synonymous with attachment loss). Additionally, the American Academy of Periodontology has discarded the awkward term probing attachment level. Furthermore, clinical attachment level is not the same as **relative attachment level**. Relative attachment level refers to the measurement from a stationary reference point such as a cusp tip, the margin of a restoration, or a stent to the most apical extent of the probe penetration, whereas clinical attachment level uses the CEJ as the reference point (Weinberg and Eskow 2003; American Academy of Periodontology 1989; Armitage 1996).

12.1.3.2 Periodontal Probe Classification

For consistency of use and academic purposes, in 1992, Pihlstrom (1992) classified probes into three generations. In 2000, Watts extended this classification by adding fourth- and fifth-generation probes. Various periodontal probes

cater to different needs. Selecting the appropriate periodontal probe is dependent on the type of dentistry a clinician practices. The needs of a general dentist are different from those of a periodontist who usually requires a more specialized set of periodontal probes. Research facilities and academic institutions can afford and effectively use more complex and sophisticated periodontal probes. Also, because the latest generations of probes work in conjunction with computers, the state of computerization in a dental practice has to be considered during the selection process (Ramachandra et al. 2011).

First-Generation (Conventional) Probes

Conventional or manual probes do not control for probing pressure and are not suited for automatic data collection. These probes most commonly are used by general dental practitioners as well as periodontists. Invented in 1936 by periodontist Charles H.M. Williams, the Williams’ periodontal probe is the prototype or benchmark for all first-generation probes. These probes have a thin stainless steel tip of 13 mm in length and a blunt tip end with a diameter of 1 mm. The graduations on these probes are 1, 2, 3, 5, 7, 8, 9, and 10 mm. (The 4 and 6-mm markings are absent to improve visibility and avoid confusion in reading the markings.) (Ramachandra et al. 2011).

The Community Periodontal Index of Treatment Need (CPITN) was designed by Professors George S. Beagrie and Jukka Ainamo in 1978. CPITN probes are recommended for use when screening and monitoring patients with the CPITN index. The index and its probes were first described in the World Health Organization’s (WHO) ‘Epidemiology, etiology, and prevention of periodontal diseases. Report of a WHO Scientific Group’. The FDI World Dental Federation/WHO Joint Working Group 1 has advised the manufacturers of CPITN probes to identify the instruments as CPITN–E (epidemiologic), which have 3.5 and 5.5-mm markings, and CPITN–C (clinical), which have 3.5, 5.5, 8.5, and 11.5-mm markings. CPITN probes have thin handles and are lightweight (5 g). The probes have a ball tip of 0.5 mm, with a black band between 3.5 and 5.5 mm, as well as black rings at 8.5 and 11.5 mm (Ramachandra et al. 2011).

University of Michigan O probes have markings at 3, 6, and 8 mm. A modification of this probe with Williams' markings also is available. University of North Carolina-15 (UNC-15) probes are colour-coded at every millimetre demarcation. They are the preferred probe in clinical research if conventional probes are required (Ramachandra et al. 2011).

The Naber's probe is used to detect and measure the involvement of furcal areas by the periodontal disease process in multirrooted teeth. Naber's probe also is used in the assessment of more complex clinical cases, including those with a restorative treatment. These probes can be colour-coded or without demarcation (Ramachandra et al. 2011).

Second-Generation (Constant-Pressure) Probes

Second-generation probes were developed in an effort to standardize and quantify the pressure used during probing. According to Hefti et al., some research 'identified a positive correlation between probing force and depth of probe penetration'. Weinberg et al. (2010) stated that controlled force of 20- to 25-g probes reduced examiner error and made depth changes of less than 2 mm clinically meaningful. The second-generation probes did not have electronic data collection. The technology of second-generation probes was the basis of the third-generation probes, which included the electronic data collection capability (Emmerling and Standley 2010).

The true pressure-sensitive (TPS) probe, introduced by Hunter in 1994, the prototype for second-generation probes, has a disposable probing head and a hemispheric probe tip with a diameter of 0.5 mm. A controlled probing pressure of 20 g is usually applied. These probes have a visual guide and a sliding scale where two indicator lines meet at a specified pressure. Other second-generation probes were designed by Armitage in 1977 (Armitage et al. 1977), van der Velden in 1978 (Van der Velden and de Vries 1978), and Polson in 1980 (Polson et al. 1980). Polson's electronic pressure-sensitive probe has a hand-piece and a control base that allows the examiner to control the probing pressure. The pressure is

increased until an audio signal indicates that the preset pressure has been reached. Polson's original design was modified by its initial users: That probe is known as the Yeaple probe, which is used in studies of dentinal hypersensitivity (Ramachandra et al. 2011).

Third-Generation (Automated) Probes

In spite of the advances in second-generation probes, other sources of errors, such as in reading the probe, recording data, and calculating attachment level, still needed to be addressed. Third-generation probes were developed to help minimize these mistakes by using not only standardized pressure but also digital readouts of the probes' readings and computer storage of data. This generation includes computer-assisted direct data capture to reduce examiner bias and allows for greater probe precision. These probes require computerization of the dental operator and can be used by periodontists and academic institutions for research (Ramachandra et al. 2011).

The Foster-Miller probe (Foster-Miller, Inc, Waltham, MA) is the prototype of third-generation probes. Devised by Jeffcoat et al. in 1986, this probe has controlled probing pressure and automated detection of the cemento-enamel junction (CEJ) (Jeffcoat et al. 1986). The components of their probe include a pneumatic cylinder that pushes the core of a linear variable differential transducer, the force transducer, the accelerometer, and the probe tip forward until they attain the final position at the bottom of the pocket. Tip movement is at controlled speed and preset force. As the probe glides along the root surface, the tip is subject to abrupt changes in acceleration when it meets the CEJ and when it is stopped at the pocket base. The tip then retracts automatically. Attachment level is computed based on the time the probe tip used to move between the two acceleration bursts and the speed of the tip movement. The probe tip consists of a Teflon-coated thin steel wire with a ball-shaped end 0.5 mm in diameter. Probing force is adjustable between 0 and 0.49 N, corresponding to probing pressures between 0 and 2.50 N/mm², respectively (Hefti 1997).

The Florida Probe System (Gibbs et al. 1988) features constant probing force, precise

electronic measurement, computerized data capturing, and sterilization of all system parts entering or close to the mouth. The system consists of a probe handpiece, a digital readout, foot switch, computer interface, and computer. The probe tip reciprocates through a sleeve and is connected to a movable arm, which transfers the movement to a transducer with digital readout. Initially, the probe tip resembled the Michigan 'O' probe tip. It was tapered and had a spherical end measuring 0.4 mm in diameter. The adjustable probing force is usually set at 0.25 N (1.99 N/mm²). The tip is flexible, a feature which improves its ability to fit the tooth contour. The slightly enlarged diameter reduces the probing pressure to 1.57 N/mm². Besides the pocket depth probe, two versions of the Florida Probe® handpiece are available for the determination of relative attachment levels: the stent probe and the disk probe. The former uses an acrylic stent as a reference and for the reproducible placement of the probe tip. The latter has a small metal disk attached to the sleeve and uses the occlusal surface or incisal edge of a tooth as a reference for relative attachment level measurement (Hefti 1997). The Florida sleeve probe handpiece was modified by Preshaw et al. (1999) to create the Florida PASHA probe, with the capability to detect the CEJ. The CEJ Probe has a modified sleeve, which includes a 0.125-mm prominent edge to facilitate a 'catch' of the CEJ. The width of this edge was considered small enough not to interfere with probing depth measurements, offering clinicians measurement of CAL and probing depth concurrently (Karpinia et al.). A clear advantage for Florida Probes is the capability of obtaining measurements to the nearest 0.1 mm, while with the conventional probe, measurements must be recorded to the nearest whole millimetre (Osborn et al. 1990).

Goodson and Kondon (1988) had used fibre-optic technology in their controlled-force **Accutek probe**. Measurement is accomplished by the movement of an optical encoder transduction element attached to a 0.35-mm-diameter disposable flexible fibre probe element. The optical encoder is attached to a fibre-optic bundle that transmits light to the transducer element and reflected light to a signal processing package. Measurement is

accomplished by counting reflected light pulses from a zero position defined initially by the probe fibre being flush with the sleeve. The electronic probe measures depths in the range of 0–12 mm to the nearest 0.5 mm under conditions of controlled probing force (Goodson and Kondon 1988). The **Interprobe™** electronic probe system is similar to the Accutek system. It is calibrated for a constant 0.3 N (1.26 N/mm²) probing force and uses a 0.55-mm-diameter plastic filament (Hefti 1997). No significant differences in probing measurements were found between Florida Probe and Interprobe (Rams and Slots 1993).

The Toronto Probe (Birek et al. 1987) is an automated periodontal probe which measures gingival attachment level using the occlusal or incisal surface of the tooth as a fixed landmark and which transfers data directly to a computer. It consists of a digital length gauge connected to a nickel–titanium alloy wire of 0.50-mm diameter enclosed in a polyethylene sheath. The wire, which serves as the probe, is propelled by air pressure into the gingival sulcus with a regulated force. Consistency of probe angulation is assured by incorporating into the probe handle a mercury level switch which will only allow data output when the instrument is held in a vertical position. Probing forces from 0.1 to 0.9 N, corresponding to probing pressures from 0.51 to 4.58 N/mm² can be generated by an electric torque motor contained in the length gauge (Hefti 1997). Similar precision data were obtained with the automated Toronto Probe compared with the pressure-sensitive probe (McCulloch et al. 1987). The Toronto Probe has also the ability of measuring probe penetration velocity (distance traversed by the probe tip/unit of time). Changes in probing velocity have been found to correlate with the degree of inflammation (Tessier et al. 1993).

Other third-generation probes are **Hunter probe** (Perry et al. 1994) and **Peri-Probe system** (Quirynen et al. 1993). The Peri-Probe system includes a constant-force probe and a data processor/printer unit. The probing force is generated by a coil spring and ranges from 0.21 N (1.07 N/mm²) for deep pockets to 0.46 N (2.34 N/mm²) for the measurement of shallow pockets. A flexible wire with a diameter of 0.3 mm and a ball-shaped

end 0.5 mm in diameter is used for probing. The probe tip is contained in a disposable sleeve, the end of which serves as a reference for probing depth measurements (Hefti 1997).

An experimental periodontal sensor probe equipped with an optical fibre for recording function was recently developed (Ishihata et al. 2012). A modified conventional periodontal probe equipped with an **optical fibre sensor** to record the probing depth was developed, and the probing depth reproducibility of the system was evaluated. The results indicated that for sites of periodontal maintenance patients with pocket depth <7 mm there was excellent agreement obtained by a single examiner using a sensor probe compared to a conventional probe. The experimental sensor probe incorporating the conventional probe tip may be useful as a routine diagnostic tool for periodontal disease, while serving as an automated data collector (Ishihata et al. 2012).

Fourth-Generation Probes

Fourth generation refers to three-dimensional (3D) probes. Currently under development, these probes are aimed at recording sequential probe positions along the gingival sulcus. They are an attempt to extend linear probing in a serial manner to take into account the continuous and 3D pocket being examined (Ramachandra et al. 2011).

Fifth-Generation Probes

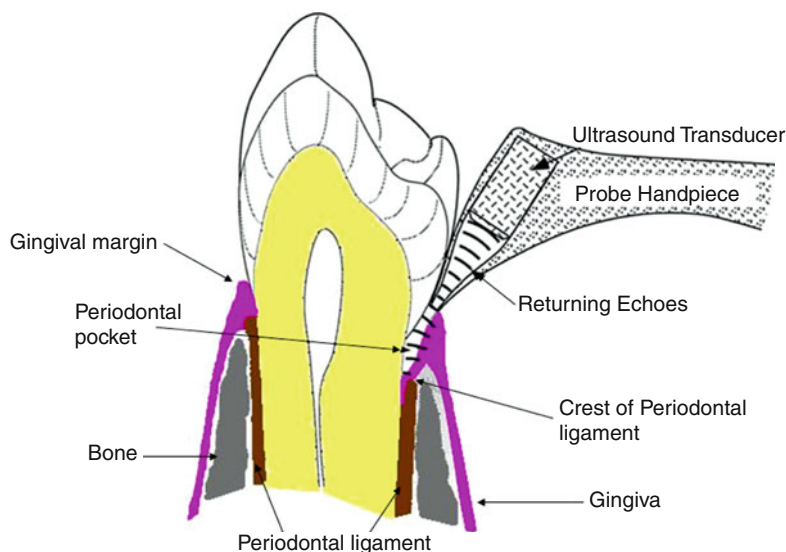
Despite all the advances in earlier generation probes, they remain invasive and, at times, their use can be painful to patients. Plus, with these earlier generation probes, the probe tip usually crosses the junctional epithelium. Fifth-generation probes are being devised to eliminate these disadvantages. Probes are being designed to be 3D and noninvasive: An ultrasound or other device is added to a fourth-generation probe. Fifth-generation probes aim to identify the attachment level without penetrating it (Ramachandra et al. 2011).

The only fifth-generation probe available, the **ultrasonographic (US) probe** (Visual Programs, Inc, www.usprobe.com), uses ultrasound waves to detect, image, and map the upper boundary of the periodontal ligament and its variation over

time as an indicator of the presence of periodontal disease. The US probe was devised by Hinders and Companion at the NASA Langley Research Center. This small intraoral probe has an ultrasound beam projection area close enough in size to the width of the periodontal ligament space to give the optimal coupling and small enough to inspect the area between the teeth, while still delivering sufficient signal strength and depth of penetration to image the periodontal ligament space. To probe these structures ultrasonically, a narrow beam of ultrasonic energy is projected down between the tooth and bone from a transducer, which is scanned manually along the gingival margin (**Fig. 12.3**). The transducer is mounted at the base of a dual-taper, convergent-divergent coupler to provide an acoustically tapered interface with a throat area on the order of 0.5 mm. This constitutes an active area reduction from the transducer element to the aperture of 20:1. Such a reduction is mandated by the geometry and the very small window afforded by the gingival margin. An added virtue of attaining this small a tip size is the ability of the ultrasonic probe to help the clinician examine the area between the teeth, which is where periodontal disease is most likely to occur (Ramachandra et al. 2011; Lynch and Hinders 2002).

The ultrasound transducer is mounted in the probe tip shell, which also incorporates a slight flow of water to ensure good coupling of the ultrasonic energy to the tissues. The couplet water can come either from a suspended intravenous-type sterile bag or plumbed from the dental-unit water source. The focused ultrasonic beam is transmitted into the pocket in the same orientation as the insertion of a manual probe. Then, the probe is moved along the gingival margin, so the two-dimensional graphical output corresponds to the results a clinician gets from 'walking the sulcus' with a manual probe. However, ultrasound gives more information because secondary echoes are recorded from tissue features at various depths. It appears likely that the technique also will be able to provide information on the condition of the gingival tissue and the quality and extent of the epithelial attachment to the tooth surface. This may supply valuable data to aid the

Fig. 12.3 Operation of the ultrasonographic periodontal probe. The probe tip is placed at the gum line with a thin ultrasound beam projecting into the periodontal pocket (Lynch et al. 2006. Reprinted with permission from Elsevier)



clinician in the diagnosis and treatment charting of these diseases (Ramachandra et al. 2011).

The ultrasonographic periodontal probe offers the potential to detect much smaller increments of anatomic change in pocket depth, thereby promising earlier detection of tissue breakdown and faster intervention. Additionally, this technology eliminates operator variability in pressure, placement, and angulation, operator errors in visual interpretation of millimetre markings, and misrecording of measurements. The ultrasonographic probe will also eliminate the need for antibiotic prophylactic premedication prior to probing in those patients who are at risk of developing bacteraemias from dental procedures. It is also expected that this procedure will be painless, be much faster than manual probing, and yield more useful patient information by providing a graphical representation of changes in pocket depth (Lynch et al. 2006).

12.1.3.3 Sources of Error for Periodontal Probing Measurements

Current probing methods are subject to various errors that were summarized by Leroy et al. (2010) as follows:

- The extent to which a probe penetrates into a given pocket:** This can vary because **inflammation** at the base of a pocket reduces resistance to a probe tip and may permit it to penetrate the base of the pocket (Robinson and Vitek 1979; Caton et al. 1981; Garnick et al. 1989)
- The diameter of the probe tip:** When the literature discussing periodontal probe diameters in human, dog, and monkey studies was reviewed and compared, it was shown that tip diameters varied from 0.4 to over 1.0 mm in these studies. Probe advancement between the gingiva and the tooth is determined by the pressure exerted on the gingival tissues and resistance from the healthy or inflamed tissue. The pressure is directly proportionate to the force on the probe and inversely proportionate to the probe tip diameter. The larger probing diameters reduced probe advancement into inflamed connective tissue. This effect of change in probe diameter reduced the pressure in a greater manner than an increase of similar change in probe force. The authors recommended that probe tips need to have a diameter of 0.6 mm and a 0.20 g force (50 N/cm²) to obtain a pressure which demonstrates approximate probing depth (Garnick and Silverstein 2000 Keagle et al. 1989).
- The tine** (part of the probe with markings): When comparing different probe tine shapes with relatively low probing forces (Atassi et al. 1992; Barendregt et al. 1996) or higher probing forces (Barendregt et al. 1996), more

shallow pockets were assessed with a tapered tine. This is most likely due to the tapered shape that gradually meets more resistance when inserted into the periodontal pocket.

- **The probing pressure** used is a further major factor influencing probe penetration: Using higher forces leads to more reproducible depth readings. On the other hand, probing with light forces allows the detection of changes, which are inaccessible by heavy probing. However, using extremely small forces may lead to underestimation of clinical attachment level gain because it primarily reflects tissue shrinkage (Mombelli and Graf 1986; Mombelli et al. 1992; Larsen et al. 2009). In a recent systematic review, Larsen et al. (2009) showed that, in the investigated studies, probing pressures range from 51 to 995 N/cm². For the evaluation of the results, a distribution was made between diseased and healthy/treated sites. The incremental change in PPD in healthy/treated sites decreased as the pressure increased above 398 N/cm². In diseased sites, this phenomenon was already present at pressures above 100 N/cm². At healthy/treated sites, a mean increase of PPD of 0.002 mm per increase of 1 N/cm² in probing pressure could be calculated, whereas at diseased sites this value amounted to 0.004 mm.
- **The angulation of the probe tine** to the pocket wall (Watts 1987, 1989): Flexible splints have been used in some studies to identify reference points and to standardize the direction of probe insertion (Watts 1987, 1989). It was shown that a variation of probe position in the transverse plane is an important source of probing error, even when a stent is used (Watts 1989).
- **The accuracy or otherwise of markings on the tine:** Although manufacturing methods to produce periodontal probes have been improved, a possible source of errors can reside in the accuracy of their millimetre markings (Winter 1979; Van der Zee et al. 1991; Neto et al. 2001).
- **The probe handle diameters:** When two periodontal probes with handle diameters of 54 and 92 mm were used, the mean overall force with the thin probe was 55.2 g, and with the thick probe, 59.4 g. However, the clinical relevance of

this difference may be minor when considering the interindividual variance of forces exerted when probing (van Weringh et al. 2006).

- **Experience of the examiner:** A study performed by Abbas et al. (1982) evaluated the effect of training upon inter-examiner agreement and concluded that only a training programme together with a standardized calibration of the probing force leads to reproducible probing depth measurements. Quirynen et al. (1993) observed that the intra-examiner deviation is smaller than the inter-examiner deviation and recommended that a single examiner should follow the same subject during longitudinal clinical trials.
- **Presence of (overhanging) restorations.**

12.1.3.4 Manual Versus Electronic Probing of the Periodontal Attachment Level

As Leroy et al. (2010) summarized, in terms of reproducibility of measurements for probing pocket depths and attachment loss, studies performed by Perry et al. (1994) and Tupta-Veselicky et al. (1994) suggested that there were no significant advantages in the use of second- or third-generation probes. This was also confirmed in a systematic review of clinical trials evaluating the reproducibility of manual and electronic probes in the measurement of clinical periodontal attachment level in untreated periodontitis subjects (Silva-Boghossian et al. 2008). **Table 12.1** presents the data as regards attachment level extracted from the two included studies (Grossi et al. 1996; Alves Rde et al. 2005). In the study of Alves Rde et al. (2005), the Pearson's correlation test showed significantly similar *r* values for CAL between manual (0.64) and electronic probes (0.57; *P* < 0.001). The absolute mean difference for CAL measurement was 0.93 (±0.86) for manual probes and 1.08 (±0.96) for electronic probes. Grossi et al. (1996) found similar average variances for replicate measurements of CAL and relative attachment level (RAL), 0.32 (0.56) and 0.33 (0.57), respectively. The authors also observed that the average variances for CAL measurements were significantly higher for molars (0.53) than anterior (0.22) and premolar

Table 12.1 Summary of extracted results from selected studies (Silva-Boghossian et al. 2008) (Reprinted with permission from Elsevier)

Study	Year	Results	
		Manual probe	Electronic probe
Alves Rde et al. (2005)	2005	Absolute mean difference of CAL measurements: 0.93 ± 0.86 ($P=0.088$) Pearson's correlation test: 0.64	Absolute mean difference of CAL measurements: 1.08 ± 0.96 ($P=0.285$) Pearson's correlation test: 0.57
Grossi et al. (1996)	1996	Average variance (standard error of measurement) for replicate measurements of CAL: 0.32 (0.56) Average variance (standard error of measurement) related to increasing pocket depth—CAL data: 0.32 (0.56) Average variance for tooth Anterior teeth: 0.22 Premolar teeth: 0.35 Molar teeth: 0.53 ($P<0.05$) Average variance for site Non-interproximal: 0.16; $P<0.05$ Interproximal: 0.41 Buccal: 0.31 Lingual: 0.34 Multiple regression analysis with the log of site specific CAL variances as the dependent variable Examiners: 0.015 Patients: 0.034 Depth: 0.064 Tooth type: 0.067 Non-interproximal: 0.075 Buccal/lingual: 0.075 PD: 7.1% CAL: 7.5% All: 7.5%	Average variance (standard error of measurement) for replicate measurements of RAL: 0.33 (0.57) Average variance (standard error of measurement) related to increasing pocket depth—RAL data: 0.33 (0.58) Average variance for tooth Anterior teeth: 0.32 Premolar teeth: 0.33 Molar teeth: 0.42 Average variance for site Non-interproximal: 0.32 Interproximal: 0.34 Buccal: 0.31; $P<0.05$ Lingual: 0.36 Multiple regression analysis with the log of site specific RAL variances as the dependent variable Examiners: 0.056 Patients: 0.109 Depth: 0.112 Tooth type: 0.113 Non-interproximal: 0.113 Buccal/lingual: 0.114 PD: 13.5% RAL: 11.4% All: 11.4%

CAL clinical attachment level, RAL relative attachment level, PD probing depth, CV coefficient of variance

(0.35) teeth ($P<0.05$) and for interproximal (0.41) than non-interproximal (0.16) sites ($P<0.05$). The average variances for RAL measurements were significantly different only for lingual (0.36) compared with buccal (0.31) sites ($P<0.05$). Using multiple regression analysis, probing depth (PD) accounted for 7.7% and CAL for 7.5% of the variance when a manual probe was used. When an electronic probe was used, 13.5% of the variance was explained by PD and 11.4% by RAL. Silva-Boghossian et al. (2008) concluded that manual and electronic probes showed a tendency to have similar reliability in the measurement of CAL in untreated periodonti-

tis subjects when used by a calibrated examiner, but note that this finding is not supported by strong evidence (Leroy et al. 2010). On the other hand, Magnusson et al. (1988) and Osborn et al. (1990) found greater reproducibility when no first-generation probes were used.

12.2 Radiographic Assessment of Periodontal Disease

Radiographic methods provide information about hard tissue changes. Radiographic images are unable to reveal soft tissue changes including

changes in periodontal attachment levels. Radiographic images compared over time indicate possible cumulative changes over a period of time, while probing values may reflect temporal changes in levels of soft tissue inflammation. Radiographs have the advantage that they can be stored and can be reassessed at any time. Radiographic assessments, however, are also subject to multiple sources of error. These include variations in projection geometry, exposure, and processing errors as well as masking of osseous structures by various anatomic structures and examiner skill in defining landmarks (Zybutz et al. 2000).

Various radiographic projections have been designed to increase the likelihood of obtaining different types of information.

Bite-wings are considered the most accurate intraoral radiographs for determining the height of the alveolar crest. In the absence of bone loss, the alveolar crest is generally located 1–2 mm apical to the CEJ. Vertical bite-wings may need to be taken to visualize the osseous crest in a patient with attachment loss. **Periapical radiographs** give the clinician important information regarding crown to root ratio, the periodontal ligament space, and the presence of periapical abnormality (Wolf and Lamster 2011). The advantages of **panoramic radiography** compared to intraoral radiography are the convenience and the simplicity of the procedure for the patient and the clinician. Panoramic radiography is also unique as it gives an overview of the entire dentition and related structures (Åkesson et al. 1992). However, panoramic radiography has a number of drawbacks which limit its usefulness as a diagnostic tool in periodontology. One of the main limitations is the potential for image distortion. The fan beam is directed upwards and a sharp image layer is formed by the coordinated movement of the beam and the receptor. Horizontal and vertical magnifications are determined by different mechanisms and vary independently as a function of the location of a structure within the image layer. Even when the patient is placed in an ideal position, lingual structures will be projected higher on the film than buccal structures. Although most units are equipped with positioning guides,

patient-positioning errors represent one of the main sources of error in panoramic radiography. This technique sensitivity also makes it difficult to reproduce the imaging geometry at a later date (Mol et al. 2004).

Radiographic quantification of crestal bone level is a valuable aid in evaluating the periodontal support of teeth (Albandar and Abbas 1986). Three methods are most often employed: (a) an absolute method which measures the distance from the alveolar bone to a reference point (usually the CEJ) (mm) (Albandar and Abbas 1986), (b) a relative method which transforms the alveolar bone height to a fraction of the radiographic root length (Schei technique) (Schei et al. 1959), and (c) a relative method employing the radiographic tooth length instead of root length (Bjorn technique) (Bjorn et al. 1969)

12.2.1 Digital Subtraction Radiography (DSR)

Using image processing, the radiographic diagnosis based upon standardized radiographs can be improved by means of a single scan densitometry, a histogram evaluation of grey-level distributions in regions of interest, corrected mean grey level, and a computer-assisted densitometric image analysis (CADIA). To quantify bone density, references in the form of aluminium wedges can be included in the images. Density changes in defined regions, such as treated periodontal sites, can then be compared to density changes in untreated sites. Image processing includes contrast enhancement and computer-aided pattern recognition. A strict chain of exclusion criteria must be applied to avoid false-positive readings. The number of non-readable pairs of images may greatly influence the accuracy of the radiographic conclusions (Brägger 2005 and references therein).

The concept of DSR is relatively simple. When two images of the same object are registered and the image intensities of corresponding pixels are subtracted, a uniform difference image will be produced. If a change in the follow-up image has occurred, this change will show up as a brighter area when the change represents gain

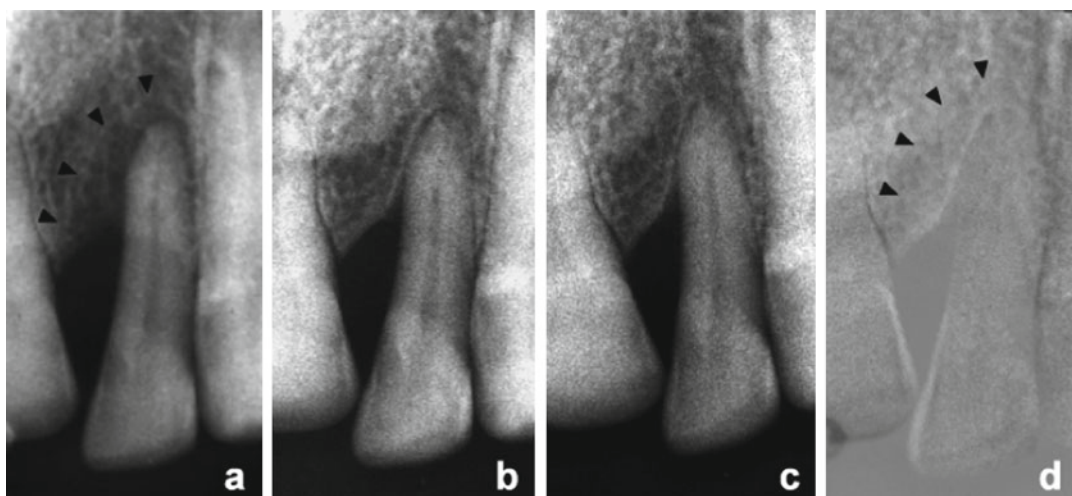


Fig. 12.4 Radiographically visible healing results in a deep intrabony defect at a maxillary left lateral incisor treated with the experimental Polydioxanon membrane: (a) intrabony defect (*arrows*) at baseline; (b) after 12 months; (c)

after 24 months; (d) digital subtraction image (24 months–baseline): *bright gray shades* (*arrows*) indicate bone density gain after 24 months (Christgau et al. 2002. Reprinted with permission from John Wiley & Sons, Inc.)

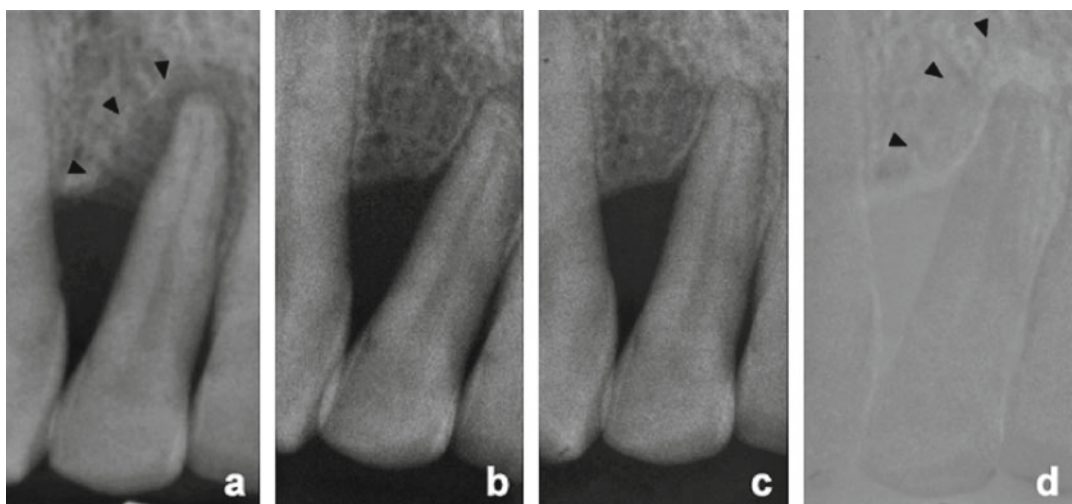


Fig. 12.5 Radiographically visible healing results in a deep intrabony defect at a maxillary right lateral incisor treated with the control Poly(lactic acid) membrane (same patient as in Fig. 12.4): (a) intrabony defect (*arrows*) at baseline; (b) after 12 months; (c) after 24 months; (d)

digital subtraction image (24 months–baseline): *bright gray shades* (*arrows*) indicate bone density gain after 24 months (Christgau et al. 2002. Reprinted with permission from John Wiley & Sons, Inc.)

and as a darker area when the change represents loss (Figs. 12.4 and 12.5) (Mol 2004).

In order for DSR to be diagnostically useful, it is imperative that the baseline projection geometry and image density be reproduced. Standardization of these image acquisition parameters is

usually considered a serious limitation, but it is precisely this requirement that makes the comparison of serial radiographs more meaningful (Mol 2004). The projection geometry is defined by the position and orientation of the X-ray source, the patient, and the detector relative to

each other. If the projection geometry used for the follow-up image is different from the projection geometry used for the baseline image, the subtraction image will show these differences. They can be difficult to distinguish from actual change within the patient or may hide actual change. Perfect reproduction of the projection geometry would be ideal, but is virtually impossible to achieve. However, not all changes in projection geometry are equally detrimental. Most changes can be reversed through image processing. There are only two changes in the projection geometry that cannot be reversed: differences in horizontal beam angulation and differences in vertical beam angulation (Mol 2004).

Subtraction images allow detection of mineral changes of as little as 5%. In addition to early detection, a number of quantitative measurements can be made, such as linear, area, perimeter, and density measurements. All can be helpful in quantifying bone mineral changes within a specific site, providing an opportunity for early assessment of the course of the disease. Various approaches for detection and quantification have been reported in the literature. While some rely upon visual interpretation and manual measurement, others use some form of automated image analysis. Image analysis algorithms designed to automatically distinguish between true osseous changes and background noise are fundamentally dependent on threshold selections and on assumptions regarding the characteristics of radiographic signs of osseous change and the characteristics of the background noise. Trade-offs between the level of automation and the generalizability of image processing applications are not uncommon. Fully automated image analysis techniques can only attain high accuracy if the radiographic signs of osseous change can be uniquely characterized. This characterization is complicated by the unpredictable behaviour of noise introduced by suboptimal image standardization. Therefore, better image standardization usually results in more accurate estimates of actual osseous change (Mol 2004).

To express change observed in subtraction images quantitatively in absolute units, mg of

bone mineral requires that X-rays be taken with a reference stepwedge (Mol 2004). The prime method employed for quantitative subtraction radiography using relative units is known by its acronym CADIA. It preselects a so-called region of interest (ROI) and, within this region as seen on the subtraction image, multiplies the area (pixels) of change by their mean grey-level change. This permits tracking of mineral change quantitatively over time. However, the CADIA technique does not permit quantitative comparison of CADIA change at different ROIs. Such comparison requires a mineral or aluminium stepwedge for internal references (Hausmann 2000).

12.2.2 Analysis of Bone Structure—Fractal Analysis

Fractal analysis is a method to analyse bone structure as reflected by change in trabecular number, size, and shape (Fig. 12.6). In this way, it differs from traditional radiographic analysis, where change is influenced primarily by mass. An advantage of fractal analysis of oral X-rays is that it is not sensitive to moderate changes of radiographic projection geometry. Therefore, standardized techniques for taking intraoral X-rays are not necessary (Hausmann 2000).

Fractal analysis was described also as a possible diagnostic indicator of osteoporosis, a significantly higher fractal dimension being observed in the group of postmenopausal (ages, 62.5 ± 4.1) compared with premenopausal (ages, 32.8 ± 3.9) women. In healthy women, the alveolar process radiographic fractal dimension was significantly related to the alveolar process density but is not related to the density of the spine, hip, or radius (Southard et al. 2001). Shrout et al. (1998) found that fractal dimensions could be used to distinguish between gingivitis and periodontitis patient groups, and fractal dimensions for gingivitis patients were reproducible over a 3-month period. Fractal analysis also evidenced significant differences between patients affected and not affected by periodontitis (Fig. 12.7) (Urdike and Nowzari 2008).

Fig. 12.6 Steps used in image processing. (a) Original dental radiograph in anterior mandible. (b) Blurred image created using a Gaussian filter with a sigma of ten pixels. (c) Image of original radiograph made uniform in overall lighting intensity by subtracting blurred image (b) from original (a). (d) Binary image, derived from Fig. 12.1. Panel c using ImageJ software (Updike and Nowzari 2008). Reprinted with permission from John Wiley & Sons, Inc.)

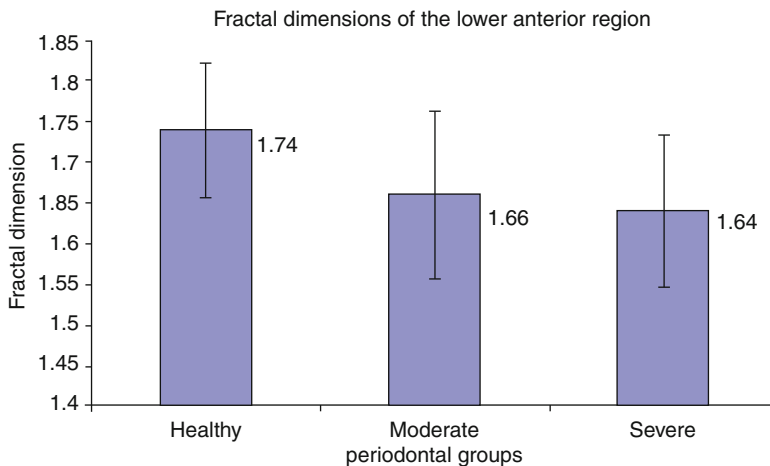
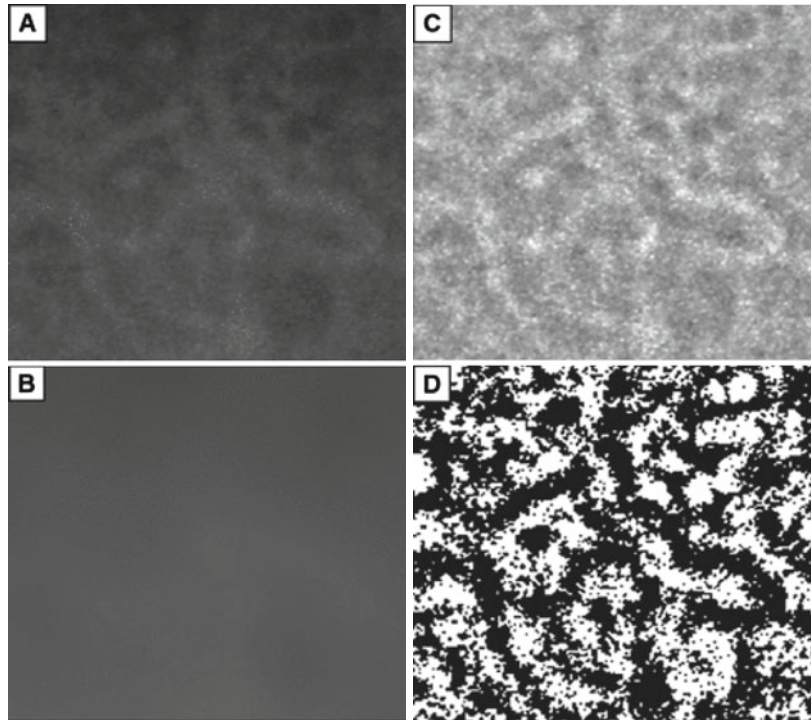


Fig. 12.7 Fractal dimensions of the three periodontal groups in the mandibular anterior area. The healthy group consisted of patients with <3 mm pocket probing depth and no other periodontal problems, such as clinical attachment loss, in the region of interest (attachment loss is defined as gingival recession plus the probing depth). The moderate periodontitis group consisted of patients with 3–4 mm of clinical attachment loss in at least 66% of the

teeth in the region of interest. The severe periodontitis group consisted of patients with ≥ 5 mm of clinical attachment loss in at least 66% of the teeth and with many involved teeth having guarded to poor prognosis in the region of interest. The error bars extend one standard deviation away from the means (Updike and Nowzari 2008). Reprinted with permission from John Wiley & Sons, Inc.)

12.3 Newly Emerging Methods for Periodontal Diagnosis

12.3.1 Cone-Beam Computed tomography (CBCT)

Numerous efforts have been made towards three-dimensional (3D) radiographic imaging (e.g. stereoscopy, tuned-aperture CT), and although CT has been available, its application in dentistry has been limited because of cost, access, and dose considerations. The introduction of cone-beam computed tomography (CBCT) specifically dedicated to imaging the maxillofacial region heralds a true paradigm shift from a 2D to a 3D approach to data acquisition and image reconstruction (Scarfe and Farman 2008).

CBCT is a recent technology. Imaging is accomplished by using a rotating gantry to which an X-ray source and detector are fixed. A divergent pyramidal- or cone-shaped source of ionizing radiation is directed through the middle of the area of interest onto an area X-ray detector on the opposite side. The X-ray source and detector rotate around a rotation fulcrum fixed within the centre of the region of interest. During the rotation, multiple (from 150 to more than 600) sequential planar projection images of the field of view (FOV) are acquired in a complete, or sometimes partial, arc. This procedure varies from a traditional medical CT, which uses a fan-shaped X-ray beam in a helical progression to acquire individual image slices of the FOV and then stacks the slices to obtain a 3D representation. Each slice requires a separate scan and separate 2D reconstruction. Because CBCT exposure incorporates the entire FOV, only one rotational sequence of the gantry is necessary to acquire enough data for image reconstruction (Scarfe and Farman 2008).

Being considerably smaller, CBCT equipment has a greatly reduced physical footprint and is approximately one-quarter to one-fifth the cost of conventional CT. CBCT provides images of highly contrasting structures and is therefore particularly well suited for the imaging of osseous structures of the craniofacial area. The use of CBCT technology in clinical dental practice provides a number of **advantages** for maxillofacial imaging (Scarfe and Farman 2008):

- Rapid scan time
- Beam limitation
- Image accuracy
- Reduced patient radiation dose
- Interactive display modes applicable to maxillofacial imaging
- Multiplanar reformation
- Ray sum or ray casting
- Three-dimensional volume rendering

While clinical applications of CBCT have expanded, current CBCT technology has **limitations** related to the 'cone-beam' projection geometry, detector sensitivity, and contrast resolution that produce images that lack the clarity and usefulness of conventional CT images. The clarity of CBCT images is affected by artefacts, noise, and poor soft tissue contrast (Scarfe and Farman 2008).

There are many 'standards' to be considered with CBCT imaging. However, the standard to which the dental profession is held, both by the public and among ourselves, transcends legal (standard-of-care) and technical (gold standard) definitions. The professional standard for CBCT is *appropriate* care: to choose CBCT imaging for each patient 'wisely' based on selection criteria derived from the best available evidence. In this expanding era of 3D imaging, the apparent urgency of adopting glittering new technology should be balanced with diligent discovery and patience (Scarfe 2011).

Several recent studies suggested that CBCT imaging has the potential to replace intraoral imaging for the assessment of periodontal architecture (Vandenberghe et al. 2007; Misch et al. 2006; Mengel et al. 2005; Noujeim et al. 2007; Loubele et al. 2008; Mol and Balasundaram 2008; Scarfe et al. 2006). However, most studies investigating the application of CBCT imaging to periodontal bone status are *in vitro*, although a few are *in vivo*, with either full-volume CBCT or limited-volume units used. Clinical studies would be helpful in supporting this conclusion (Tyndall and Rathore 2008).

In a recent review of currently published literature on CBCT for periodontology, Kasaj and Willershausen (2007) revealed that with this device, three-dimensional sectional images in the axial, frontal, and sagittal plane can be obtained at one examination with tomographic slices of

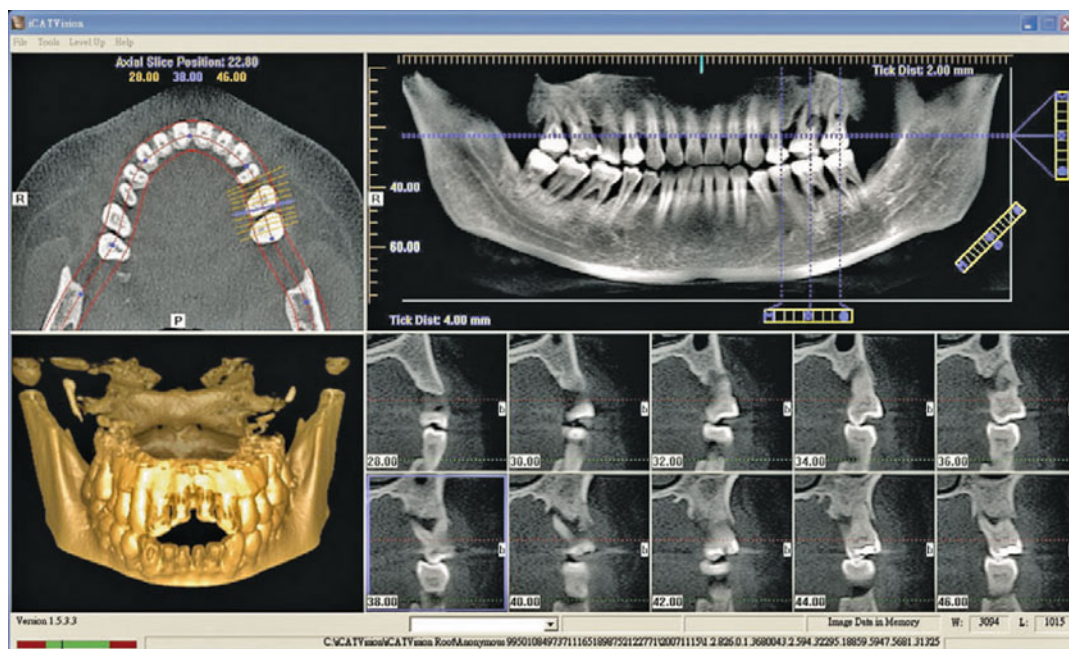


Fig. 12.8 i-CAT Vision software interface, consisting of pan-map (*upper right*), horizontal section (*upper left*), vertical sections (*lower right*) and reconstructed 3-dimen-

sional model (*lower left*) (Corbet et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

widths ranging from 0.125 to 2 mm. The authors concluded that a field of interest for the use in periodontology might be the imaging of periodontal intra-bony defects, dehiscence and fenestration defects, and periodontal cysts as well as the diagnosis of furcation-involved molars. CBCT may be a useful and more practical clinical digital subtraction radiography for the assessment of changes in periodontal bone over time (Tyndall and Rathore 2008).

Figure 12.8 shows the software interface of one type of cone-beam CT management software (the i-CAT Vision, Imaging Sciences International, Hatfield, PA, USA). The interface consists of four windows. The upper left window is the horizontal view of the arch and the lower right window is a series of the vertical views of a bony defect area. The lower left window is the three-dimensional stimulated reconstruction model of the region of interest. Operators of the viewing software can adjust the focal trough in the horizontal view shown in the upper left image, along the form of the arch in this window so that the pan-map image on the upper right window can be

constructed according to this trough. The horizontal and vertical location bars can be adjusted to centre on the defect of interest in the other two windows. **Figures 12.9** and **12.10** are the zoomed-in images of the periodontal defects centred in **Fig. 12.8**. **Figures 12.9** and **12.10** show how the three-dimensional appreciation of the defect and the tooth-roots can be elaborated. Operators can adjust the position of the image, power of zoom in and zoom out, and contrast of the images in order to assess a particular region of interest. The thickness of the slices across an area of interest can also be adjusted down to an interval of 0.2 mm (Corbet et al. 2009).

12.3.2 Optical Coherence Tomography (OCT)

OCT is a new medical imaging modality with resolution in the μm range (average, 4 μm) and depth of imaging in the mm range that can give near-histologic images with a safe broadband light source (Brezinsky 2006; Drexler and Fujimoto

2008; Ali and Parlapalli 2010). Broadband laser light waves are emitted from a source and directed towards a beam splitter. One wave is sent towards



Fig. 12.9 Horizontal views of periodontal alveolar bony defects at upper left second premolar (25) to upper left second molar (27) (Corbet et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

a reference mirror with a known path length and the other towards the tissue sample. After the two beams reflect off the reference mirror and the tissue surfaces at varying depths in the sample, the reflected light is directed back towards the beam splitter, where the waves are recombined and read with a photo detector (Fig. 12.11). The basic principle of OCT is analogous to CT (which uses X-rays), magnetic resonance imaging (which uses spin resonance), and B-scan ultrasound (which uses sound waves) (Baek et al. 2009).

OCT can be used to image various aspects of biological tissues. Some of these include (Ali and Parlapalli 2010):

- Structural information: This is the most typical OCT imaging. It measures the local reflectivity of the tissue.
- Blood flow: Using Doppler technique, the blood flow in tissues and vessels can be estimated.
- Polarization sensitive: The change of polarization states through tissues can be estimated.
- Elastography: The elastic parameters of the tissue can be estimated.
- Spectroscopy: The variation of absorption, reflectivity, and scattering with wavelength can be measured, providing clues to molecular content of the tissue.

Any combination of the above imaging modes can be used to bring out specific features of biological tissues as desired.

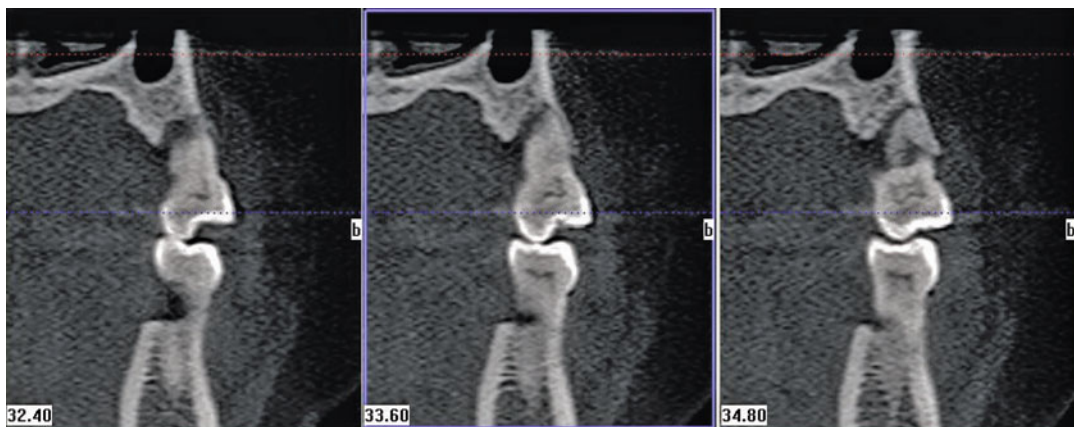


Fig. 12.10 Vertical views showing lingual furcation involvement at lower left first molar (36) and an extensive defect at palatal aspect of the upper left first molar (26)

(Corbet et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

Applications of OCT in various biomedical imaging applications include (Ali and Parlapalli 2010):

- Ophthalmology: Very fine imaging of the retina with capability to identify several eye diseases is possible with OCT. Most of the commercial products available today are for ophthalmologic usage.
- Dermatology: To image subsurface structural and blood flow information.
- Dentistry: To image structure of teeth and gum line at the same time to visualize bacteria in concert with the tooth and roots.
- Gastroenterology: To image the gastrointestinal (GI) tract through endoscopic probes.

- Intra-vascular: To image plaques inside blood vessels.
- Cancer diagnosis: Several modes in OCT imaging can discriminate between malignant and normal tissues, allowing cancer diagnosis through either noninvasive or minimally invasive procedures.
- Intra surgery for tumour margining: Enables discrimination between malignant and non-malignant tissue to decide the regions of tissue to be removed during surgery.

A study performed by Otis et al. (2000) suggested that *in vivo* dental OCT images clearly depict anatomical structures that are important in the diagnostic evaluation of both hard and soft oral tissue. Periodontal tissue contour, sulcular depth, and connective tissue attachment are visualized at high resolution using this technology (Figs. 12.12 and 12.13). Because OCT reveals microstructural detail of the periodontal soft tissues, it offers the potential for identifying active periodontal disease before significant alveolar bone loss occurs. OCT images of the periodontium can be stored in the patient record, providing visual documentation of disease progression, response to therapy, or both. Similar data were presented by Xiang et al. (2010a) who showed that OCT imaging can offer three-dimensional imaging of periodontal soft tissues and bone at an exquisitely high resolution. Because OCT reveals microstructural details of the periodontal soft tissues, it offers the potential for identifying active periodontitis before significant alveolar bone loss occurs (Xiang et al. 2010a).

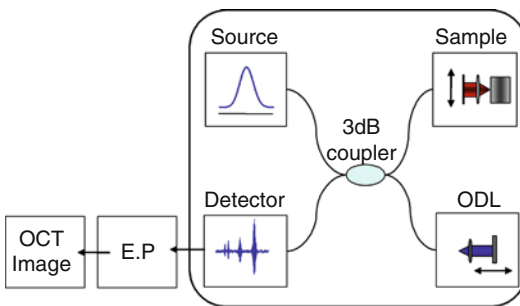


Fig. 12.11 Schematic image of OCT system: *ODL* optical delay line, *E.P.* electrical processing. At first, the light source is split into the sample and the ODL by using the coupler. Then the detector detects interferences between signals from the ODL and the sample when reflected light from the sample and the ODL are recombined. Finally, electrical processing changes these interferences into visible optical images (Baek et al. 2009. Reprinted with permission from Elsevier)

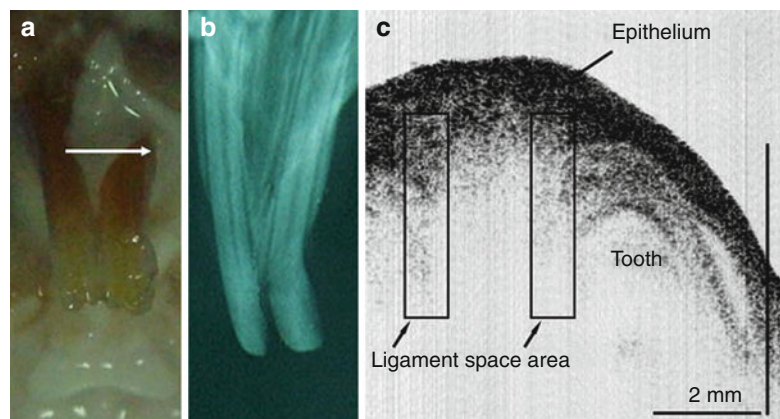


Fig. 12.12 Maxillary anterior tooth of a white rat in the control group, to which no orthodontic force was applied: (a) optical photograph; (b) radiographic image; (c) OCT image (Na et al. 2008. Reprinted with permission from Springer Science + Business Media)

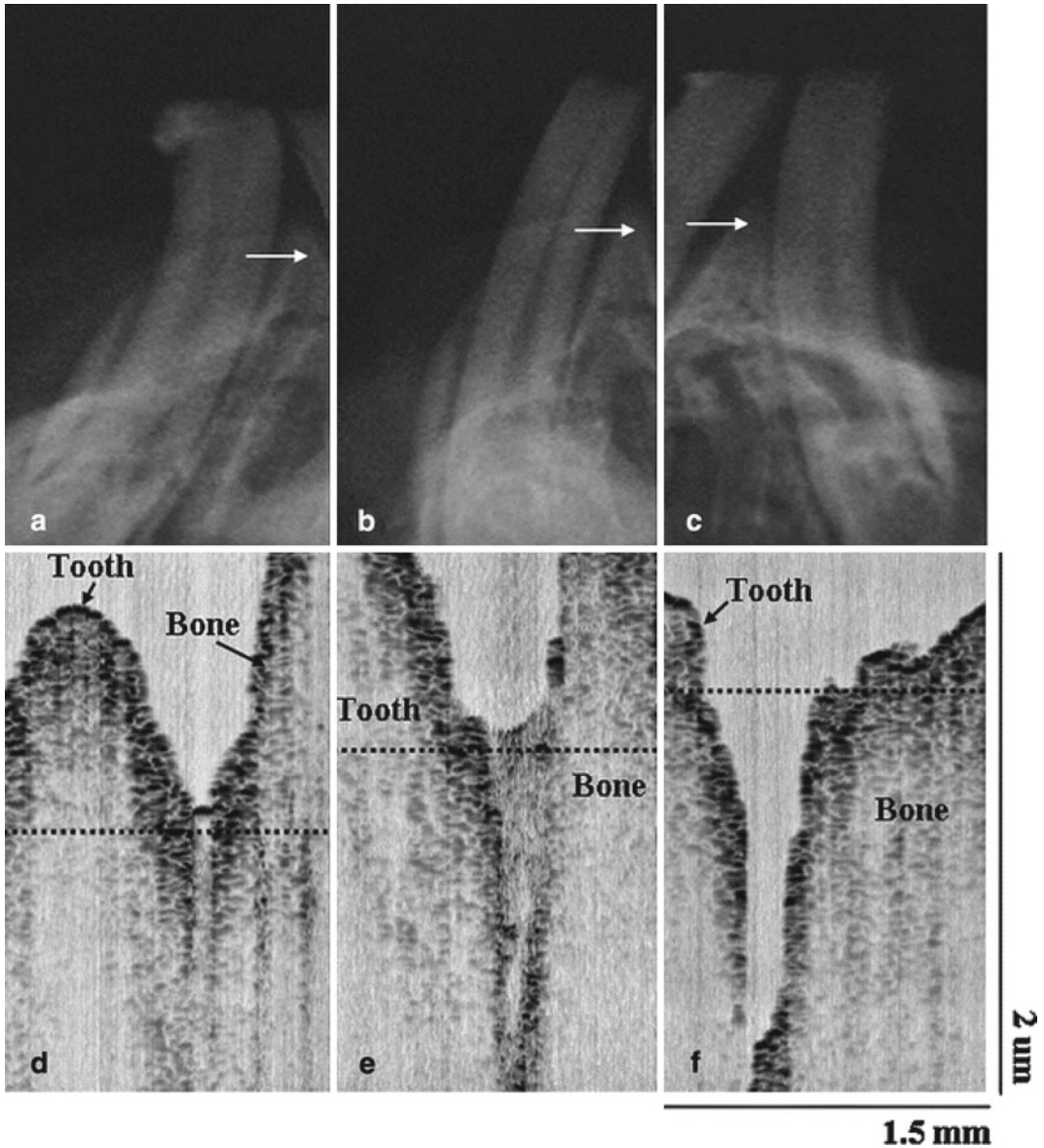


Fig. 12.13 Radiographic (*top*) and OCT images (*bottom*) of the maxillary anterior teeth of white rats under several orthodontic forces: (a, d) 5 gf; (b, e) 10 gf; (c, f) 30 gf forces, respectively. The two-dimensional OCT image in

the *bottom* was taken along the arrow direction in each of the *top* (radiographic) figures (Na et al. 2008. Reprinted with permission from Springer Science + Business Media)

OCT systems are signal processing intensive. The need for acquiring real-time data, processing the acquired data to extract meaningful information, and then displaying the information in a clinically relevant way is well suited for embedded implementations using digital signal processors and system-on-chip application processors.

The advent of low-power digital signal processors and system-on-chip is an added benefit for OEMs (*original equipment manufacturers*) to develop low-cost, low-power, and even portable systems based on OCT technologies without compromising the image quality needed for clinical applications. In addition, the programmability

and scalability of these devices allow optimizing the signal chains for different applications on the same platform (Ali and Parlapalli 2010).

12.3.3 Tuned-Aperture Computed Tomography (TACT)

Imaging systems used in dentistry, including the field of periodontology, are largely limited to 2-dimensional (2D) systems. Two-dimensional systems include conventional film-based radiography and digital radiography with the use of a digital sensor as an image capture mechanism. The problem inherent to 2D systems is that three-dimensional (3D) anatomy is collapsed into 2D space, resulting in the superimposition of structures that potentially obscure features of interest and decrease diagnostic sensitivity. Therefore, there continues to be a need for radiographic tools with a 3D perspective for pretreatment and post-treatment assessment of periodontal defects (Ramesh et al. 2001).

There are a number of 3D systems available, but they all have limitations. Conventional tomography exhibits blurring due to superimposition of out-of-focus structures. Computed tomography is constrained by its high cost and limited spatial resolution relative to transmissive radiographic systems (Ramesh et al. 2001).

Tuned-aperture computed tomography (TACT) is a 3D image-forming algorithm that could be implemented with virtually any projection-based imaging system capable of digitized output. It produces true 3D data from any number of arbitrarily oriented two-dimensional (2D) projections. It differs from existing modalities in that it computes the 3D representation using information acquired from known fiducial relationships in 2D projections. This is in contrast to the conventional method, wherein reconstruction is accomplished from 'back projections' relative to a fixed-projection geometry designed into the associated hardware (Nair et al. 2002).

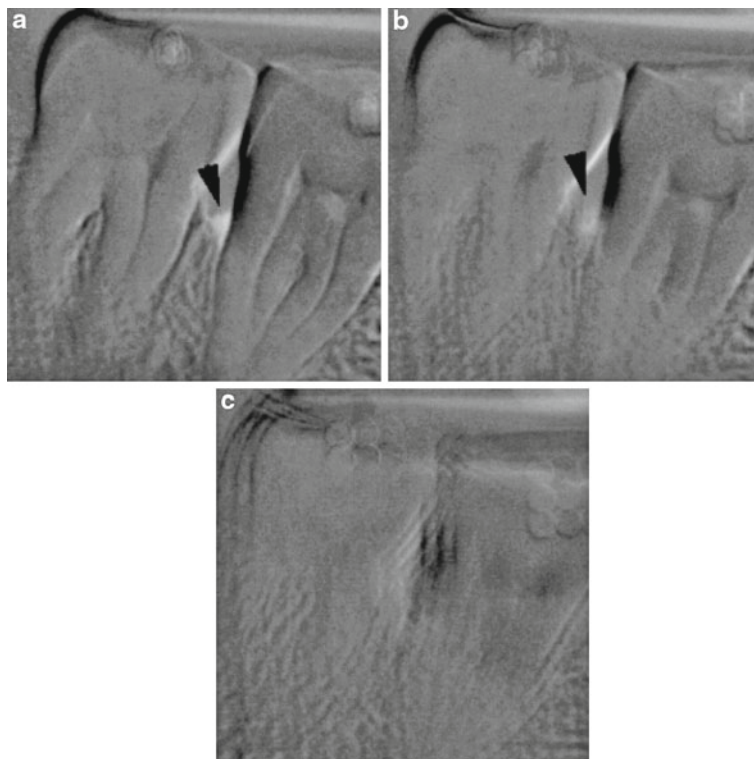
Studies have been done to assess performance of TACT for various diagnostic tasks in dentistry and concluded that TACT has a potential for improved diagnostic performance for primary

caries (Tyndall et al. 1997; Harase et al. 2006), simulated recurrent caries (Mjor et al. 1997; Nair et al. 1998), osseous jaw lesions (Horton et al. 1996; Webber et al. 1996; Nair et al. 2007), impacted third molars (Harase et al. 2005), implants (Horton et al. 1996; Webber et al. 1996; Rashedi et al. 2003; Limrachtamorn et al. 2004), endodontics (Barton et al. 2003; Nair et al. 2003a), and periodontal defects (Ramesh et al. 2001, 2002; Nair et al. 2001; Nair and Bezik 2006). The incorporation of TACT with digital subtraction radiography analysis (TACT-DSR) frees the process of longitudinal image acquisition from the rigorous constraint of image geometry, thus possibly overcoming the main limitation to the clinical application of DSR. TACT-DSR has a great potential to provide greater sensitivity and technique flexibility in detecting periodontal bone-gain than standard DSR (Fig. 12.14) (Chal-U-Dom et al. 2002).

In direct comparison with TACT, cone-beam computed tomography (CBCT) is limited in resolution. TACT can use images from any type of image sensor such as film or solid-state devices. Therefore, the resolution of the resulting images is the same as that of the original images. Resolution in the range of 16–24 lp (line pairs)/mm can thus be achieved, in addition to the generation of images with higher signal-to-noise ratios. CBCT resolution is significantly inferior to that of TACT as sensor pixel sizes are relatively larger. Furthermore, the amount of scatter produced in volumetric tomography modalities such as CBCT is higher, too, in comparison with TACT. The potential for artefact generation in the presence of metallic restorations exists. However, CBCT provides a better alternative to conventional medical grade CT for a variety of routine maxillofacial diagnostic tasks at radiation doses that are much lower and at a reduced cost. The other advantage of TACT is that no dedicated equipment is necessary for image acquisition and reconstruction. However, at this point, it is a time-consuming task. In the future, the process could be automated. The selection of appropriate imaging modalities based on the diagnostic task assumes importance in this context as each modality is different (Nair and Bezik 2006).

Fig. 12.14 Representative TACT-DSR slices for the area of the lower right second and third molars with 4 mg pericrestal bone-gain distal to tooth #31. Displays are slices in buccal (a), middle (b) and lingual sites (c). Notice the area of brighter gray level in (a) and (b) but not (c); this demonstrates an area of bone-gain in the buccal portion of the defect (Chal-U-Dom et al. 2002.

Reprinted with permission from John Wiley & Sons, Inc.)



12.3.4 Magnetic Resonance Imaging (MRI)

Nuclear magnetic resonance, also called MRI, has been available as an imaging technique since 1984. MRI combines the use of a magnetic field and some radio frequency antennas called coils. During the MRI exam, a magnetic field is created: This causes the protons in the atoms of water, within the tissues to be examined, to line up. Then pulses of radio waves are sent from a scanner towards the tissue. A high-frequency pulse moves many of the protons out of alignment. As the nuclei realign back into proper position, they send out the radio signals (resonance) that are captured by a radio antenna. The signals are received and measured from a computer system that analyses and converts them into an image of the part of the body being explored. Different atoms in the body absorb radio waves at different frequencies under the influence of the magnetic field. The way in which absorption takes place is

measured and used by the computer to construct the images (Cotti and Campisi 2004).

Magnetic resonance is a completely noninvasive technique since it uses radio waves; it also allows acquisition of direct views of the body in almost all orientations. Its best performance is in showing soft tissues and vessels, whereas it does not provide great details of the bony structures. The strength of the MRI system magnetic field is measured in metric units called Tesla. In MRI, bright means high signal intensity, dark means low signal intensity, with all the intermediate shades, that is, dense bone=dark, air = dark, and fat=bright. Using a special programme, STIR (short tau inversion recovery, fat annulling sequences), water, and blood will appear bright. The pictures of MRI will look, as in CT, like sections or cuts. Disadvantages of MRI are a longer scanning time compared to CT and a strong magnetic field generated, which is why it cannot be used in patients carrying a pacemaker or metal pieces in the areas to be investigated. It is an

expensive examination, and in most of the systems, the patient must be placed in a narrow tube (Cotti and Campisi 2004).

In medicine, MRI is used to distinguish pathological tissues from normal tissues in regions where other imaging modalities do not provide sufficient spatial resolution and contrast (Kwock et al. 2002). In the orofacial region, MRI is used mainly for diagnosing the pathology of the temporomandibular joint (Rammelsberg et al. 2000), the floor of the mouth (Barbiera et al. 2002), the tongue (Barbiera et al. 2002), salivary glands (Takagi et al. 2005), and paranasal sinuses (Eggesbo et al. 2001). Due to its excellent soft tissue contrast and its high sensitivity to detect oedema, magnetic resonance imaging (MRI) might represent a complementary imaging technique to visualize particular pathological processes of the jaw and teeth apparatus, for example, teeth pulp pathologies, periapical lesions, inflammatory disease of the periodontal space, or teeth vascularization after trauma. On the other hand, MRI of the dental apparatus is technically challenging as the bony structures of the mandible and teeth give only little or no MR signal (Gaudino et al. 2011).

MRI was used in diagnosis of pulp pathosis on extracted teeth, being revealed that with this technology pulp tissue can be imaged in spite of its location within calcified tissue (Lockhart et al. 1992). When the anatomic structure of rat teeth was studied and observed using magnetic resonance imaging with high spatial resolution, the periodontal ligament dividing the teeth from the bony structures were easily observed (Fig. 12.15a, b). The pulp and ligament surrounding the incisor were seen in hypersignal. Details of the crown are shown in Fig. 12.15c. All the anatomic structures of the rat mandible were clearly visualized (Beuf et al. 1997).

The diagnostic feasibility of MRI for demonstrating granulocytic sarcoma occurring in the maxillary gingiva (Lee et al. 2001), for the visualization of periodontal tissues in porcine mandible (Gaudino et al. 2011), or for the characterization of the inflammation and healing processes in periodontal tissues in patients with moderate to advanced periodontal disease (Schara et al. 2009) was also investigated.

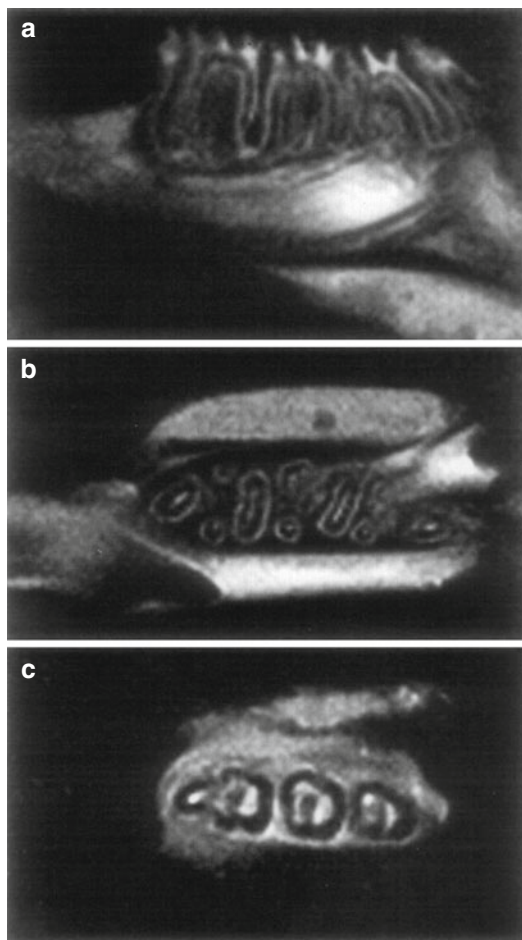
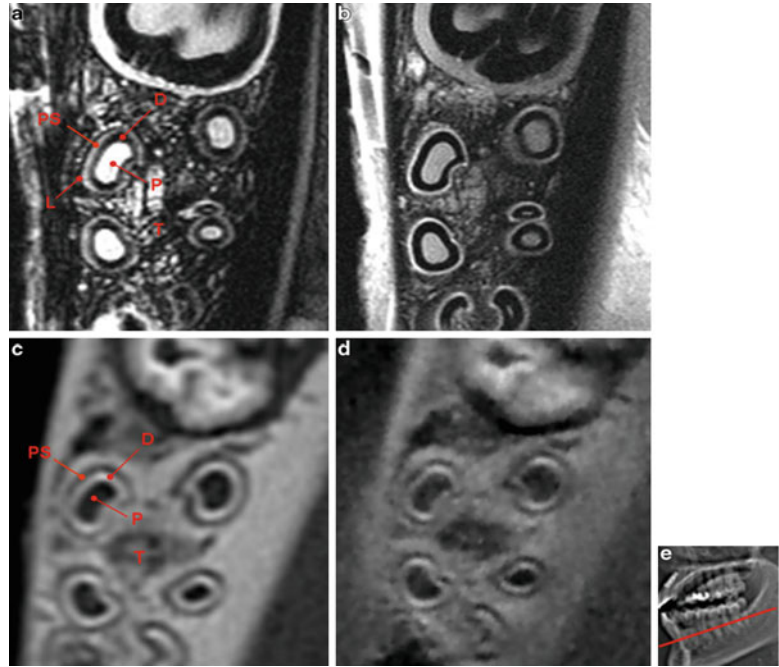


Fig. 12.15 High-resolution magnetic resonance image ($1.05 \times 1.5 \text{ cm}^2$) of the rat mandible in an adult subject with a $120 \times 120 \text{ g}^2$ in-plane resolution. (a) Sagittal view (TR=3,000 ms, TE=40 ms) and coronal views (TR=1,500 ms, TE=23 ms) obtained by (b) slice selection at the molars roots level and (c) at the crown (Beuf et al. 1997. Reprinted with permission from Elsevier)

When the ability of magnetic resonance imaging (MRI) to visualize dental and periodontal structures was assessed and findings with multi-detector computed tomography (MDCT) and cone-beam CT (CBCT) were compared, it was shown that periodontal space and cortical/trabecular bone were better visualized by MRI ($P<0.001$). MRI could excellently display the lamina dura, not detectable with MDCT and only inconsistently visible with CBCT ($P<0.001$) (Fig. 12.16) (Gaudino et al. 2011).

Fig. 12.16 Depiction of anatomical structures of the dental roots and periodontal apparatus on axial T2 (a) and T1 weighted (b) MR images, MDCT (c) and CBCT (d). Image plane is parallel to the mandible major axis and perpendicular to the roots, as shown schematically on a human localizer in (e). *P* pulp chamber, *D* dentin, *PS* periodontal space, *L* lamina dura, *T* trabecular bone. Notice that the lamina dura is not visible on MDCT and CBCT (Gaudino et al. 2011. Reprinted with permission from Springer Science + Business Media)



The results of MRI provided also a new possibility to characterize the type and healing process of periodontal inflammation, as after the therapy the decrease of inflammation in periodontal tissues and the improvement in clinical parameters were confirmed by signal intensity measurements of MR images (Schara et al. 2009).

12.3.5 *In Vivo* Monitoring and Screening of Periodontitis by Near-Infrared (NIR) Spectroscopy

The visible near-infrared (NIR) spectral region of the electromagnetic spectrum covers the range from 400 to 2,500 nm and conveys information on local tissue haemodynamic changes of the disease, such as oedema and oxygenation. For instance, the short-wavelength region, 500–600 nm, is dominated by the absorption from oxygenated haemoglobin (HbO₂) and deoxygenated haemoglobin (Hb) in the capillary bed of gingival tissue, whereas the absorption from water results in an increased attenuation at longer wavelengths in the 900–1,100 nm region (Ge et al. 2011).

Monitoring the intensity of the water absorption bands in the reflectance spectrum of gingival tissues provides an index of tissue hydration, representing a simple indicator of inflammation at specific periodontal sites, and can be readily detected at sub-clinical levels. The increase in blood supply is often insufficient to meet the oxygen demand in inflamed gingivae. Therefore, in addition to measuring tissue water content, NIR spectroscopy was successfully used to assess the balance between tissue oxygen delivery and oxygen use in periodontal tissues by fitting optical attenuation spectra to the known optical properties (extinction coefficients) of HbO₂ and Hb (Ge et al. 2011).

Acquisition of optical spectra: For spectra collection, a portable PDA512-ISA spectrograph system (Control Development Inc., South Bend, IN, USA) with a customized bifurcated fibre-optic bundle, or 'probe', designed for use in the oral cavity was used (Fiberguide Industries, Stirling, NJ, USA) (Fig. 12.17). The intraoral probe consisted of a 180-mm-long stainless steel shaft, 5 mm in diameter, housing 29 fibre-optic bundles. The central fibre area, which delivers light to the oral cavity, is 0.8 mm in diameter. The outer ring of fibres is ~0.2 mm wide and located

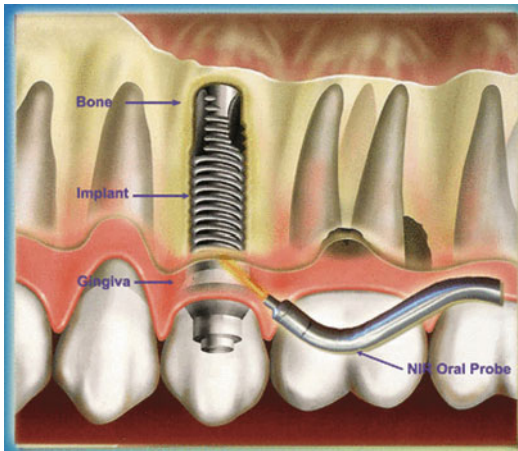


Fig. 12.17 Illustration of proposed near-infrared (NIR) oral probe as a noninvasive diagnostic tool for the assessment of peri-implantitis in the clinical setting (Nogueira-Filho et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

1 mm away from the central fibre area. These outer fibres are coupled to the entrance slit of the spectrograph and collect light subsequently emitted from the tissue. The emitter fibres at the bifurcated end of the bundle were coupled to a 5-W tungsten halogen light source (Spectral Products, Putnam, CT, USA) that provides a highly stable light output. Each reflectance spectrum consisted of 16 co-added scans collected using a 0.03-s integration time. The spectral range between 500 and 1,100 nm at 5-nm resolution was used. A 99% Spectralon® reflectance standard (LabSphere, North Sutton, NH, USA) can be used as a reference to convert raw data into reflectance spectra (Liu et al. 2009).

To calculate the clinically relevant parameters of O_2 saturation, blood volume, and hydration from the measured attenuation spectrum, the known absorption coefficients of the major chromophores present in the tissue that give rise to the spectral features observed in the spectrum are used, as described below. A modified Beer–Lambert unmixing model that incorporates a nonparametric scattering loss function is used to determine the relative contribution of deoxygenated haemoglobin, oxygenated haemoglobin, and H_2O to the overall spectrum. For the visible region between 510 and 620 nm, where deoxygenated

haemoglobin and oxygenated haemoglobin dominate the spectrum, the measured attenuation spectrum, A_λ , is modelled as the sum of two parametric terms—deoxygenated haemoglobin and oxygenated haemoglobin, which contribute to the spectrum—and a nonparametric term, $m(\lambda)$, modelling a vector of covariates that have an unspecified functional form on the composite spectrum (Liu et al. 2009).

$$A(\lambda) = \sum_{i=1}^3 \xi_i(\lambda) c_i L + m(\lambda) + \text{error}$$

Deoxygenated haemoglobin and oxygenated haemoglobin concentrations per unit photon path length are recovered by using the noniterative method for estimation in partially linear models based on kernel smoothing. The combined spectral signatures of oxygenated haemoglobin and deoxygenated haemoglobin measure total haemoglobin, an indicator of tissue perfusion, whereas the ratio of oxygenated haemoglobin to total haemoglobin represents the oxygen saturation of the tissue. Therefore, tissue oxygen saturation and tissue perfusion (total haemoglobin) are derived from the predicted deoxygenated haemoglobin and oxygenated haemoglobin concentrations, where (Liu et al. 2009)

$$S_{tO_2} = \frac{[HbO_2]}{[HbO_2] + [Hb]}$$

and

$$tHb = [HbO_2] + [Hb]$$

Early studies in the 1990s have shown that the TRS can be used for the estimation of gingival microvascular function using a dog gingiva (Hanioka et al. 1989) and human gingiva (Hanioka et al. 1990). Furthermore, an increase in haemoglobin concentration and a decrease in oxygen saturation in inflamed gingiva during induction of experimental periodontitis in dogs were found (Hanioka et al. 1989). Recently, the ability of NIR to determine simultaneously multiple inflammatory indices (tissue oxygenation, total tissue haemoglobin, deoxyhaemoglobin, oxygenated haemoglobin, and tissue oedema) in

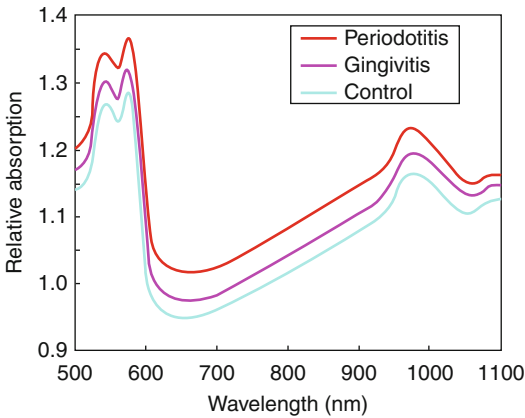


Fig. 12.18 Mean near-infrared spectra from healthy, gingivitis and periodontitis sites (Liu et al. 2009). Reprinted with permission from John Wiley & Sons, Inc.)

periodontal tissues *in vivo* was evaluated. The mean area normalized spectra generated from three groups of sites: healthy, with gingivitis, and with periodontitis, are presented in **Fig. 12.18**. The short-wavelength region, 500–600 nm, was dominated by the absorption from oxygenated haemoglobin and deoxygenated haemoglobin in the capillary bed gingival tissue, whereas the absorption from water resulted in an increased attenuation at longer wavelengths in the 900–1,100 nm region. The spectral readings between the gingivitis and healthy sites were virtually indistinguishable; however, the mean spectrum from the periodontitis group showed subtle deviations in both the haemoglobin and water regions of the spectrum. **Figure 12.19** presents the relative concentrations of deoxygenated haemoglobin and oxygenated haemoglobin extracted from the visible region of the optical reflectance spectrum from periodontal tissues *in vivo* for the three groups of sites. Additional inflammatory indices derived from these measures, such as tissue oxygen saturation and total haemoglobin (which is closely related to tissue blood volume), are presented in **Fig. 12.20** (Liu et al. 2009). It has been shown that tissue oxygenation at periodontitis sites was significantly decreased ($P < 0.05$) compared to sites with gingivitis and healthy controls. This was largely the result of an increase in deoxyhaemoglobin in the periodontitis sites compared with healthy ($P < 0.01$) and gingivitis

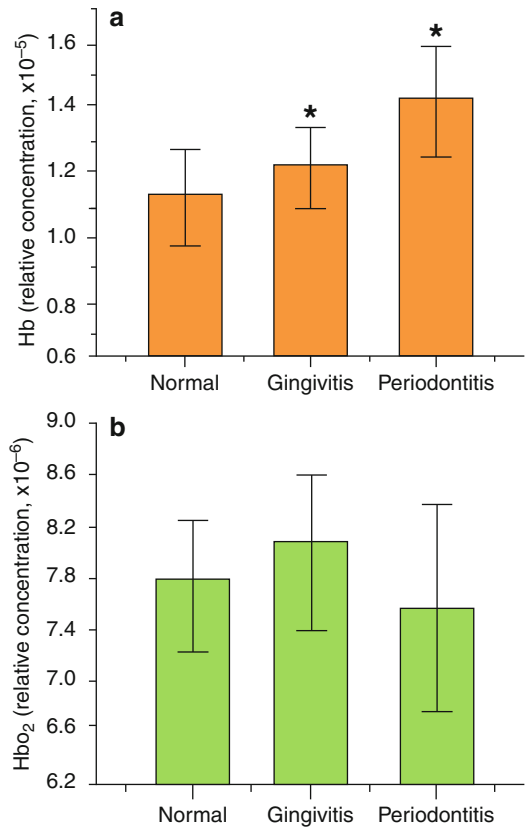


Fig. 12.19 Relative concentrations of deoxygenated hemoglobin (a) and oxygenated hemoglobin from healthy, gingivitis and periodontitis sites (b). Relative hemoglobin concentrations were calculated using the visible region (510–620 nm) of the reflected light spectrum. *Represents a significant difference from healthy sites, $P < 0.01$. Vertical bars denote 0.95 confidence intervals. Hb deoxygenated haemoglobin, HbO₂ oxygenated haemoglobin (Liu et al. 2009). Reprinted with permission from John Wiley & Sons, Inc.)

($P = 0.05$) sites. Tissue water content per se showed no significant difference between the sites, but a water index associated with tissue electrolyte levels and temperature differed significantly between periodontitis sites and both healthy and gingivitis sites ($P < 0.03$) (Liu et al. 2009). These findings were supported by a consecutive study at a geographically distinct location (Suzhou, China) to a broader patient population who also showed that tissue oxygenation as measured by optical spectroscopy is significantly decreased in periodontitis and that optical spectroscopy can simultaneously

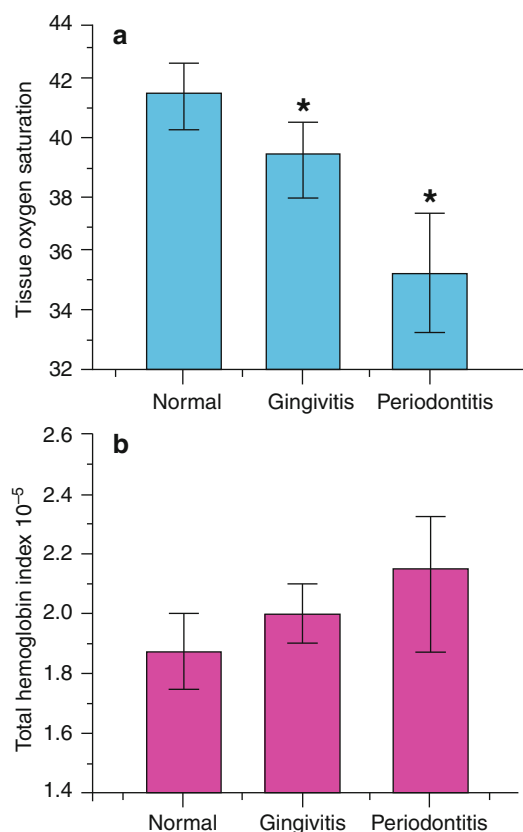


Fig. 12.20 Per cent tissue hemoglobin oxygen saturation (a) and total hemoglobin indices (b) derived from the relative concentrations of deoxygenated hemoglobin and oxygenated hemoglobin. Indices are compared among healthy, gingivitis and periodontitis sites. *Represents a significant difference from healthy sites, $P < 0.01$. Vertical bars denote 0.95 confidence intervals (Liu et al. 2009). Reprinted with permission from John Wiley & Sons, Inc.)

determine multiple inflammatory indices related to periodontal disease directly in gingival tissues *in vivo* (Ge et al. 2011)

The diagnostic potential of optical spectroscopy was further evaluated in peri-implant inflammation *in vivo* by assessing multiple inflammatory parameters (tissue oxygenation, total tissue haemoglobin, deoxyhaemoglobin, oxygenated haemoglobin and tissue oedema) simultaneously (Nogueira-Filho et al. 2011). The most striking difference was the oxygen saturation of the tissue presented in Fig. 12.21. There was a significantly lower oxygen saturation in sites diagnosed as peri-implantitis when com-

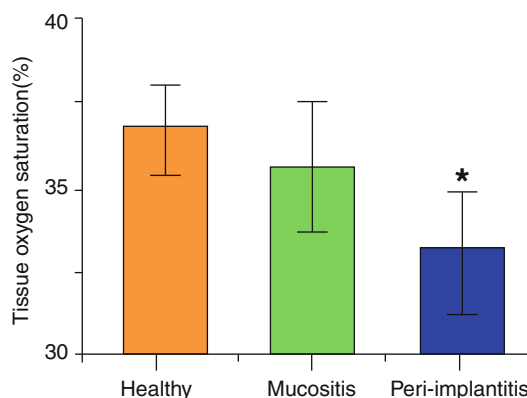
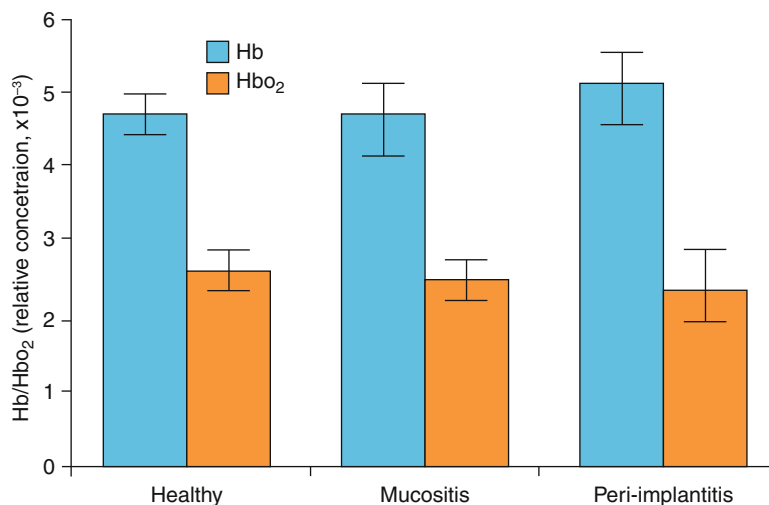


Fig. 12.21 The percentage of tissue hemoglobin oxygen saturation derived from the relative concentrations of Hb and HbO₂. Indices are compared between healthy, mucositis and peri-implantitis sites. * Significant difference from healthy sites, $P < 0.05$. Vertical bars denote 0.95 confidence intervals (Nogueira-Filho et al. 2011). Reprinted with permission from John Wiley & Sons, Inc.)

pared with healthy sites ($P < 0.05$). As mentioned in the material and methods section, the tissue haemoglobin oxygen saturation is derived from the predicted Hb and HbO₂ relative concentrations. Figure 12.22 presents the mean ($\pm 95\%$ confidence interval) relative concentrations of Hb and HbO₂ parameters obtained from the optical attenuation spectra measured *in vivo* from healthy, mucositis, and peri-implantitis sites. Although not statistically significant, there was a trend towards a higher relative concentration of Hb and a lower relative concentration of HbO₂ in peri-implantitis sites ($P > 0.05$). Moreover, the tissue hydration index was higher in diseased sites than in healthy sites. In particular, the increased tissue hydration in mucositis sites was significant when compared with both healthy ($P < 0.05$) and peri-implantitis sites, whereas no significant difference was seen between healthy and peri-implantitis sites (Nogueira-Filho et al. 2011).

NIR spectroscopy was also used to compare oxygen saturation of haemoglobin in the gingiva (GSO₂) of smokers and nonsmokers and to evaluate the chronic effect of smoking on gingival oxygen sufficiency. GSO₂ in healthy gingiva was significantly lower in smokers than nonsmokers ($P = 0.001$), while smokers showed higher GSO₂

Fig. 12.22 Relative concentrations of Hb and HbO₂ from healthy, mucositis and peri-implantitis sites. Relative hemoglobin concentrations were calculated using the visible region (510–620 nm) of the reflected light spectrum (Nogueira-Filho et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)



than nonsmokers in moderately inflamed gingiva ($P=0.035$). The GSO₂ in inflamed gingiva was significantly decreased compared with healthy gingiva in nonsmokers ($P=0.004$), while smokers showed no significant difference between healthy and inflamed gingiva ($P=0.277$ – 0.866). GSO₂ in smokers was consistently and significantly lower than that of healthy gingiva of nonsmokers ($P=0.039$ – 0.0004). These data suggest that smokers exhibit possibly lower function of oxygen sufficiency in healthy gingiva and reduced ability to adapt the function in inflamed gingiva than nonsmokers and have functional impairments in the gingival microcirculation (Hanioka et al. 2000a). It was also showed that pocket oxygen tension (pO₂), a major environmental determinant of the subgingival microflora, is lower in smokers than in nonsmokers and that the pO₂ in smokers is not influenced by gingival oxygen sufficiency, providing the basis of understanding environmental factors possibly associated with microbial flora in the pockets of smokers (Hanioka et al. 2000b).

As neither tissue oxygen saturation nor subtle local oedema is readily detectable clinically, these data suggest that optical spectroscopy could be a useful supplementary tool for diagnosis and monitoring of inflammation around dental implants. In principle, optical spectroscopy is an entirely noninvasive technique that uses low-energy radiation, spectra can be captured instantly, no consumables need to be purchased or devel-

oped, and minimal training is required to obtain reliable and reproducible data. This leads us to believe that optical spectroscopy can be an important tool in the near future, to help determine healthy and diseased dental implants, as well as to help distinguish between progressing lesions and lesions that represent the consequences of a previous disease (Nogueira-Filho et al. 2011).

The results of the above-described proof-of-concept studies reveal that haemodynamic alterations can be detected around diseased periodontitis and peri-implant sites by optical spectroscopy, offering a feasible approach for the monitoring and diagnosis of peri-implant diseases. Prospective longitudinal clinical studies are warranted to demonstrate the role of this tool in differentiating progressive from stable lesions and in monitoring the effect of periodontal and peri-implant therapies (Nogueira-Filho et al. 2011).

12.3.6 Noninvasive Diagnosis of Periodontitis by Acoustic Microscopy

Ultrasound imaging is another imaging modality that has been explored for the diagnosis of periodontitis. The application of an ultrasound as a diagnostic tool is well documented. The definition of an ultrasound is based on mechanical vibrations at frequencies above the limit of human audibility. Most diagnostically useful sound frequencies

occur between 1 and 5 MHz. Ultrasound is reflected when it strikes an interface between two media of different acoustic properties. This effect is used for the measurement of distance (Xiang et al. 2010b). The ultrasound method has as its main advantage the absence of the unwanted side effects of radiation, which is the case during radiological examination or computed tomography. The measurements could be repeated at small time intervals, obtaining reproducible data without negative effects on the patient's health (Chifor et al. 2011).

Ultrasound imaging has been extensively used in medical and dental areas such as diagnostic imaging and physiotherapy. For instance, dental scaling, a procedure for removing dental plaque and calculus from the surface of the teeth, is a very popular routine practice in dental offices. Other applications of ultrasound in dentistry include the treatment of joint- and muscle-related ailments around the face, the generalized cleaning of instruments prior to sterilization and dentures, and as a diagnostic procedure to detect dental caries and periodontal disease (Xiang et al. 2010b).

Löst et al. (1989) assessed whether the ultrasonic procedure is suitable for the determination of the height of the facial/oral bone crest. *In vitro* ultrasonic scans of the oral/facial periodontium of pigs were obtained using a special experimental design and with the help of a newly developed ultrasonic measuring and analysis system. These two-dimensional images distinguish themselves from the one-dimensional RF-echograms and A-scans in that they enable the viewer to have a quick overview of the relationship between different interfaces in a scanned sector. A selection of ultrasonic B-scans proves that the scanned tissues such as gingiva, bone, periodontal ligament, tooth surface, and even structures of the inner part of the tooth may be clearly imaged and distinctly differentiated. Furthermore, it was demonstrated with manipulations of pig's periodontium that the interpretation of the pictures is correct.

Recently, Mahmoud et al. (2010) investigated the feasibility of using a custom-designed high-frequency ultrasound imaging system to reconstruct high-resolution (<50 μm) three-dimensional (3D) surface images of periodontal defects in human jawbone (Fig. 12.23). The system

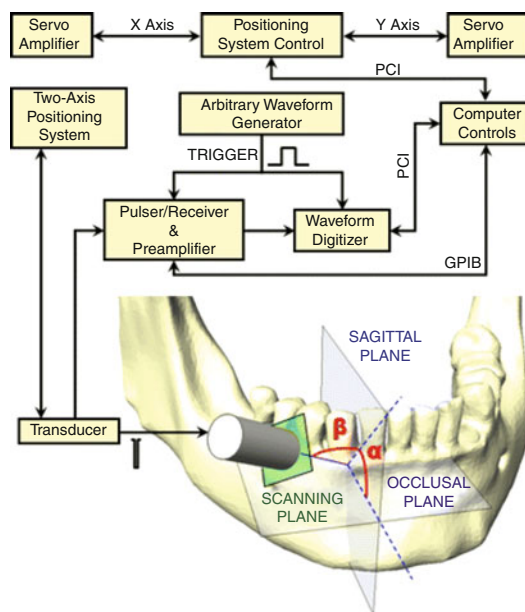


Fig. 12.23 Schematic diagram showing the design and processing steps of the high-frequency ultrasound system for jawbone (Mahmoud et al. 2010. Reprinted with permission from Springer Science + Business Media)

employed single-element focused ultrasound transducers with centre frequencies ranging from 30 to 60 MHz. Continuous acquisition using a 1-GHz data acquisition card was synchronized with a high-precision two-dimensional (2D) positioning system of $\pm 1\text{-}\mu\text{m}$ resolution for acquiring accurate measurements of the mandible, *in vitro*. Signal and image processing algorithms were applied to reconstruct high-resolution ultrasound images and extract the jawbone surface in each frame. Then, all edges are combined and smoothed in order to render a 3D surface image of the jawbone. *In vitro* experiments were performed to assess the system performance using mandibles with teeth (dentate) or without (nondentate). The system was able to reconstruct 3D images for the mandible's outer surface with superior spatial resolution down to 24 μm and to perform the whole scanning in <30 s. Major anatomical landmarks on the images were confirmed with the anatomical structures on the mandibles. All the anatomical landmarks were detected and fully described as 3D images using this novel ultrasound imaging technique, whereas the 2D X-ray radiographic images suffered from poor contrast (Figs. 12.24

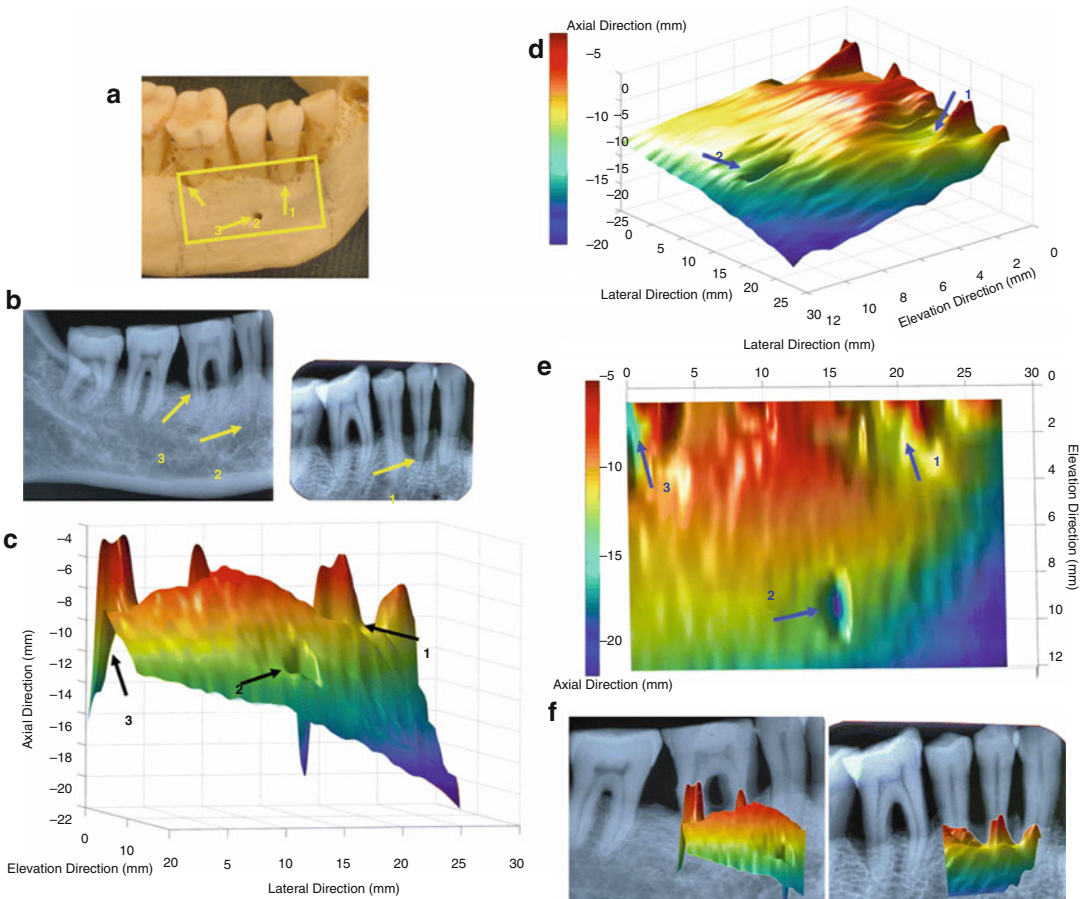


Fig. 12.24 Defected dentate dried cadaver mandible. (a) Photographic image for the mandible with a rectangle showing the scanned region and three landmarks, (b) the corresponding X-ray radiographic images showing landmarks #1, 2, and 3, (c–e) different views for the 3D ultrasound jawbone surface image. It is clearly observed that the X-ray images have very poor information about land-

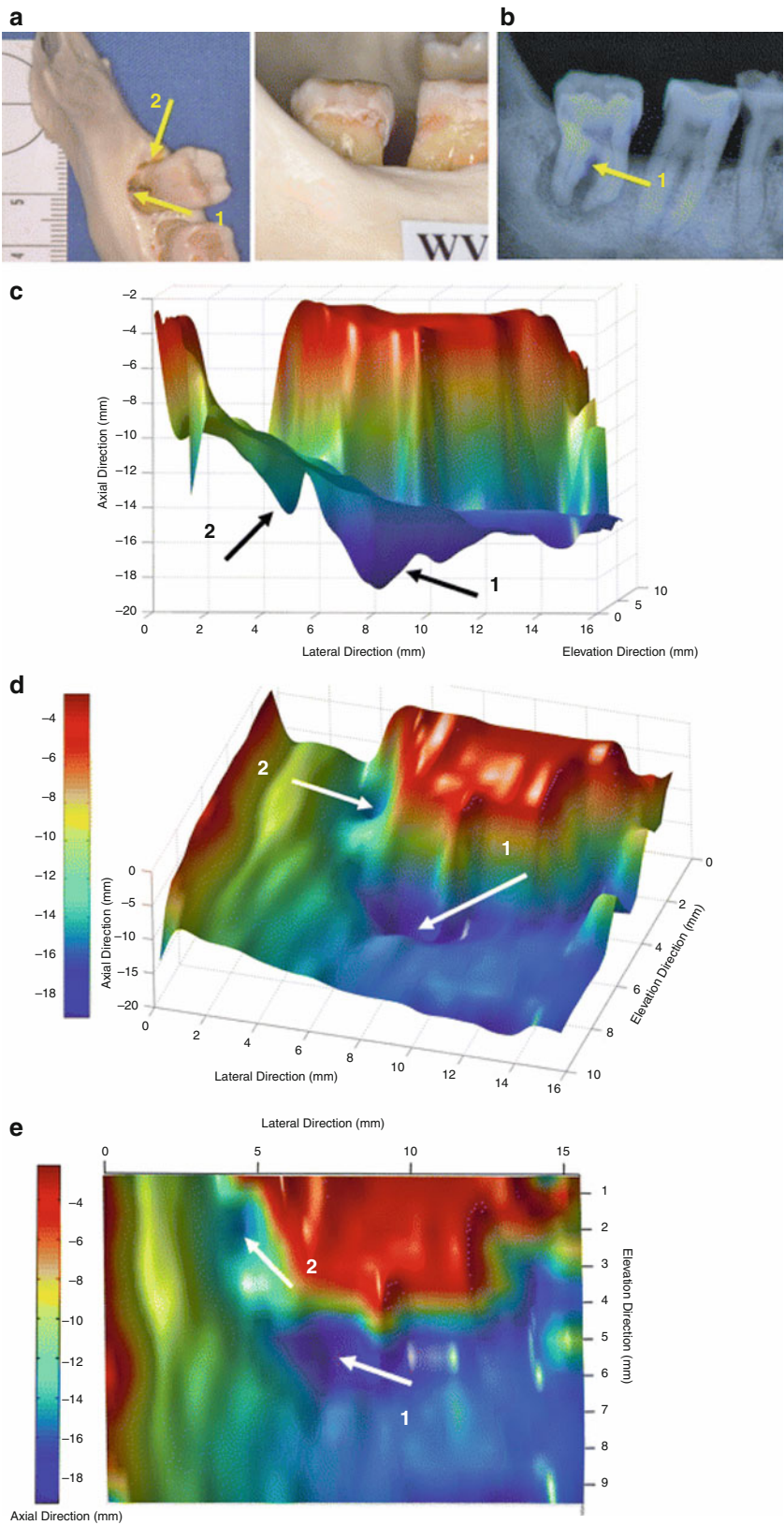
marks #2, and #3, and limited information to perfectly describe landmark #1, and (f) 3D ultrasound images approximately overlaid on the corresponding X-ray images showing featured landmarks for comparison (Mahmoud et al. 2010. Reprinted with permission from Springer Science + Business Media)

and 12.25). These results indicate the great potential of utilizing high-resolution ultrasound as a noninvasive, nonionizing imaging technique for the early diagnosis of the more severe form of periodontal disease (Mahmoud et al. 2010).

More recently, Chifor et al. (2011) identified by ultrasonography (DermaScan C, Cortex Technology) the reference points necessary to monitor the horizontal bone resorption and to assess the accuracy of the measurements by comparing with cone-beam computed tomography images, having direct microscopic section measurements as a gold standard. It was concluded that compared to the measurements of the areas

from the coronary limit of the alveolar processes performed on radiological images, the ultrasonographic data will bring equal information. On ultrasonographic images, the measurements could be done with a precision of micrometers unlike with computed tomography, where the precision is within a tenth of millimetres. The resolution is higher for ultrasonography. In the same respect, the information given by the ultrasonographic images is much fuller for buccal and lingual sides of the teeth than the information given by retroalveolar X-rays, where the cortical bone is projected over the root and it cannot be clearly seen, due to the superimposition.

Fig. 12.25 Defected dentate dried cadaver mandible. **(a)** Two photographic images for the mandible with *arrows* showing two landmarks around the third molar, **(b)** the corresponding X-ray radiographic image with an *arrow* describing landmark #1, and **(c–e)** different views for the 3D ultrasound jawbone surface image. The 3D ultrasound images were able to detect different bony defects which could not be defined using X-ray images at landmark #1 (sever defect), and landmark #2 (early stage defect) (Mahmoud et al. 2010. Reprinted with permission from Springer Science + Business Media)



It was also concluded that the cemento-enamel junction may be identified with high accuracy according to the tooth anatomical convexities. The high resolution of the images demonstrates these convexities with a very good precision, therefore enabling the user to easily identify the cemento-enamel junction. This is a very important step because previous studies (Löst et al. 1989; Mahmoud et al. 2010) positioned the alveolar bone level compared to landmarks drilled into the enamel, which may certainly not be done on real clinical cases because it is out of the question from an ethical point of view to drill cavities into healthy enamel just to get reference points for further measurements (Chifor et al. 2011).

The ultrasound method has as its main advantage the absence of the unwanted side effects of radiation, which is the case during radiological examination or computed tomography performed for that area. The measurements could be repeated at small time intervals, obtaining reproducible data without negative effects on the patient's health (Chifor et al. 2011).

This method will have a high applicability and utility in the follow-up of bone resorptions due to forces applied during orthodontic treatments, giving the possibility of follow-up and monitoring the effects of these forces on bone remodelling in the coronary area of alveolar processes (Chifor et al. 2011). Moreover, in endodontics, this method has provided very important data for the assessment of periradicular lesions of endodontic origin (Cotti et al. 2002, 2003, 2006; Cotti 2010).

In summary, ultrasound imaging can visualize periodontal and oral tissues *in vivo* or *ex vivo* without the need for complicated processing, fixing, or staining. It is fast and noninvasive. Therefore, echography is an easy and reproducible technique that has the potential to supplement conventional radiography in the diagnosis and follow-up of periodontal-associated disease (Xiang et al. 2010a).

12.4 Advances in Microbiologic Analysis

Bacterial species colonizing the surfaces of the human oral cavity are known to play an important role in oral health and disease, and thus, a

rapid and accurate means of identification is crucial. Traditionally, identification has been based on phenotypic and biochemical criteria, including microscopy, biochemical reactivity, growth conditions, dye and immunofluorescence staining, bacterial end-product analysis, cell-membrane composition, and antibiotic sensitivity. However, these tests are labour intensive and costly, providing sometimes inconsistent results that make identification rather tentative. This can be caused by strain variation within a species. More recently, molecular DNA-based techniques have been used to identify bacteria directly from clinical samples, circumventing the need for *in vitro* culture. The results from these studies implicate specific bacterial species or complexes of species that are associated with oral health and disease. In general, there are three main categories of molecular microbial analyses to consider, namely, (1) polymerase chain reaction (PCR)-based methods, including single target PCR, multiplex PCR, and quantitative PCR; (2) DNA-DNA hybridization methods, such as *in situ* hybridization, checkerboard hybridization, and 16S ribosomal RNA-based microarrays; and (3) sequencing methods including the latest, next-generation sequencing techniques, such as pyrosequencing, real-time single-molecule DNA sequencing, and nanopore-based sequencing (Paster and Dewhirst 2009)

12.4.1 Microbiological Sampling Methods

The information generated by microbiological analysis of plaque collected from a periodontally diseased site is highly dependent on the sampling technique. There are two primary methods by which a patient's subgingival plaque can be collected for subsequent analysis: removal using curettes or adsorption onto endodontic paper points. Both require careful removal of supragingival plaque at the site prior to sampling in order not to contaminate or dilute the subgingival sample. Depending on the nature of the microbiological test, samples of plaque are either analysed immediately, as in the case of chairside tests, or placed in vials containing transportation fluid and

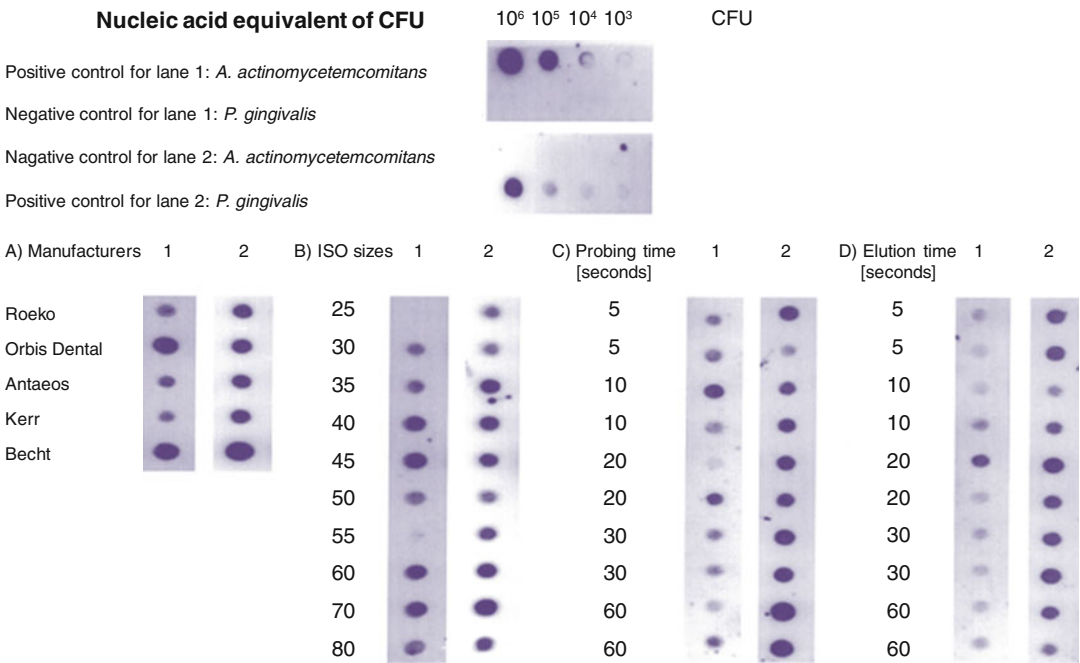


Fig. 12.26 Performance of paper points in the collection of bacterial samples for molecular genetic diagnosis: comparison of different manufacturers (Blot A), different ISO sizes (Blot B), different probing (sampling) times (Blot C), and different elution (vortexing) times (Blot D). The test specimens are: lane 1 *A. actinomycetemcomitans* and lane 2 *Porphyromonas gingivalis* (Hartroth et al. 1999. Reprinted with permission from John Wiley & Sons, Inc.)

sent to licensed clinical laboratories for analysis (Loomer 2004).

A curette or paper point is generally used to collect subgingival plaque samples. **Paper points**, which are basically introduced to dry root canals in endodontic treatment, are being more and more frequently recommended for sampling subgingival plaque for microbiological diagnostics. This method is not invasive, and the complete tips can be transferred to the transport vial after probing. They are provided with molecular genetic-based commercial detection systems for periodontal pathogens (Hartroth et al. 1999). However, paper points have certain disadvantages; if they become wet by gingival crevicular fluid they lose stiffness and it becomes difficult to move them to the bottom of the pocket. Another disadvantage is that different paper points placed in the same pocket may not sample the same microorganisms (Krigar et al. 2007). Also one possible disadvantage could be that only a selective aliquot, mainly the flowing or loosely adherent plaque, is collected by paper points and that, as a consequence, bacteria of the more tightly

attached plaque could be underestimated (Hartroth et al. 1999).

Hartroth et al. (1999) demonstrated significant differences in the performance of paper points from different manufacturers/suppliers. Maximum bacterial recovery using paper points of ISO size 45 or 70 was found. Since the latter are too thick for appropriate periodontal pocket sampling in most cases, the ISO 45 size seems to be optimum. The specification of the paper points used as given by some authors (e.g. ‘medium paper points’) may not be exact enough to describe the performance of the paper points and, as a result, of the diagnostic system used. A sampling time of 60 s seems to be optimum. Probing times longer than 60 s are not recommended, since saturation of the paper points with plaque fluid could be observed between 45 and 60 s after insertion into the periodontal pocket. Since the consistency of natural subgingival plaque differs from the pure bacterial suspension that is used in tests, probing times under 10 s also cannot be recommended in practice. Five and 30 s did not significantly reduce the sampling efficiency (**Fig. 12.26**) (Hartroth et al. 1999).

According to the joint statement of the German Society for Periodontology (DGP) and the German Society for Dental, Oral, and Maxillofacial Medicine (DGZMK), microbiological testing prior to anti-infective therapy is indicated for the following clinical diagnoses: aggressive periodontitis, generalized severe chronic periodontitis, periodontitis exhibiting progressive attachment loss despite thorough treatment, and severe periodontitis associated with systemic diseases (e.g. human immunodeficiency virus [HIV] infection). Subgingival plaque samples should be collected from the deepest pockets exhibiting signs of activity, that is, bleeding or suppuration. A microbiological analysis that is representative of the subgingival microflora of the whole oral cavity is relevant for adjunctive systemic antibiotic therapy of certain forms of periodontitis. Also because of economic reasons, the analysis of pooled plaque sampled from several sites is recommended (Beikler et al. 2005).

Subgingival plaque also may be sampled using curettes. **Curettes** are made of steel and stay stiff when wet. Further, a curette samples subgingival plaque from a larger area than a paper point. Thus, sampling plaque with a curette might overcome some of the disadvantages of paper points. However, using paper points of standardized size provides standardized samples that are preferred by the laboratory (Krigar et al. 2007).

Okada et al. (2000, 2001) have developed a new method of plaque sampling from patients by use of a **toothbrush**, which takes approximately 1 min to perform. Their report suggests that it is easy to use and does not cause any anxiety in the children, while the average amount of genomic DNA collected is sufficient for carrying out PCR.

12.4.1.1 Comparisons Between Currently Used Sampling Devices

Comparisons between currently used sampling devices can be made based on their relative ability to access defined zones in a pocket, sample size, and sample composition when the same site is resampled by the same or different devices (Tanner and Goodson 1986). However, the com-

parison of different sampling methods from the literature is complicated by a lack of standardization of basic parameters associated with the sampling performance itself, such as devices used for sampling, sampling time, or laboratory methodology (Jervøe-Storm et al. 2007).

Regarding **sample size** collected, Tanner and Goodson (1986) demonstrated that curettes collected 61–91% of the total microbiota compared to 7–41% retrieved by paper points. The reason for this difference, in part, was that paper points became saturated by loosely attached plaque; therefore, the more tightly adherent plaque was not sampled. This property of paper points was demonstrated clearly in a study by Baker et al. (1991). Paper points were placed in layered bacterial suspensions of *A. actinomycetemcomitans* followed by *A. naeslundii* and then in the reverse order, as well as in mixed suspensions of the two species. The bacteria in the top layer accounted for $\geq 90\%$ of the total colony-forming units retrieved from the paper points. Thus, paper point samples are likely to underestimate the presence of certain groups of bacteria. In another clinical study (Renvert et al. 1992), paper point samples were compared to scaler samples before and after treatment. This study concluded that paper points collected more colony-forming units and spirochetes before and after therapy than curettes. Recently, Jervøe-Storm et al. (2007) compared the two widely used sampling techniques for subgingival bacteria using quantitative real-time polymerase chain reaction. The analysis enabled the quantification of *A. actinomycetemcomitans*, *Fusobacterium nucleatum*, *Porphyromonas gingivalis*, *Prevotella intermedia*, *Treponema denticola*, and *Tannerella forsythensis* and total bacterial counts. It was shown that although curette samples harvested significantly more total bacteria, the composition of the plaque samples with respect to selected target pathogens was quite similar for both sampling techniques. Thus, both techniques can be recommended for clinical use for the microbiological testing of periodontal lesions.

When the two sampling devices were compared based on their relative ability **to access defined zones in a pocket and sample composi-**

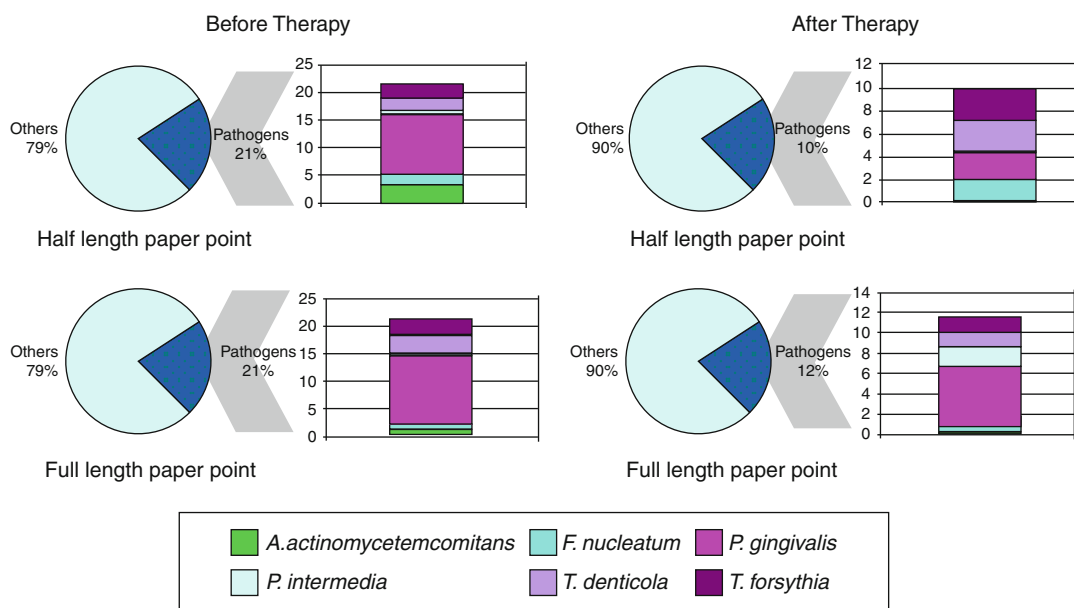


Fig. 12.27 Means of proportions of half length paper point samples and full length paper point samples in group A (Jervøe-Storm et al. 2010. Reprinted with permission from Springer Science+Business Media)

tion, early research found that each method may provide different microbiologic samples. Loomer (2004) reported that paper point samples differed from curette samples and that a curette collects plaque from the entire pocket (Renvert et al. 1992; Strand et al. 1987). In contrast, paper points collect plaque from the outer layer of the plaque, which may contain more pathogens. At the same time, paper points are less successful at sampling the apical part of the pocket, where more pathogens are expected to be. This means that the most coronal and outer portion of the plaque is sampled with paper points (Mombelli et al. 1989; Tanner and Goodson (1986)). It was also suggested that paper points may not reach the very apical portions of the periodontal pocket (Hartroth et al. 1999) and may be more susceptible than curettes to contamination with supragingival biofilm during insertion and removal from the sulcus (Smola et al. 2003). Arguably, curettes are more prone to variations in the amount of plaque collected at each sampling because difficulties in standardizing the procedure have been reported (Sixou et al. 1991). Moreover, Kornman (1986) stated that curette sampling should not be used if the microbial

ecosystem is not to be disturbed because curettes could alter the ecosystem much more than other commonly used techniques.

However, recently, when the recovery of six periodontal pathogens by paper point samples from two different aspects of periodontal lesions by quantitative real-time polymerase chain reaction (PCR) was recently compared, it was shown that the recovery of target pathogens was similar following sampling at various depths of the periodontal lesion (Jervøe-Storm et al. 2010). Twenty patients with untreated chronic periodontitis were randomized into two groups. Before subgingival instrumentation and after 10 weeks, samples in group A were taken first with a paper point half length of the probing depth, then with a paper point full length at the same site. In group B, sampling sequence was reversed. Analysis by real-time PCR enabled quantification of six bacteria as well as total bacterial count. Group A paper point samples contained the same total target pathogen proportions (TTPP) regardless of length before therapy (both lengths, 21%). After therapy, slightly higher TTPP were found with full-length paper points (half length, 10%; full length, 12%; **Fig. 12.27**). Comparing the

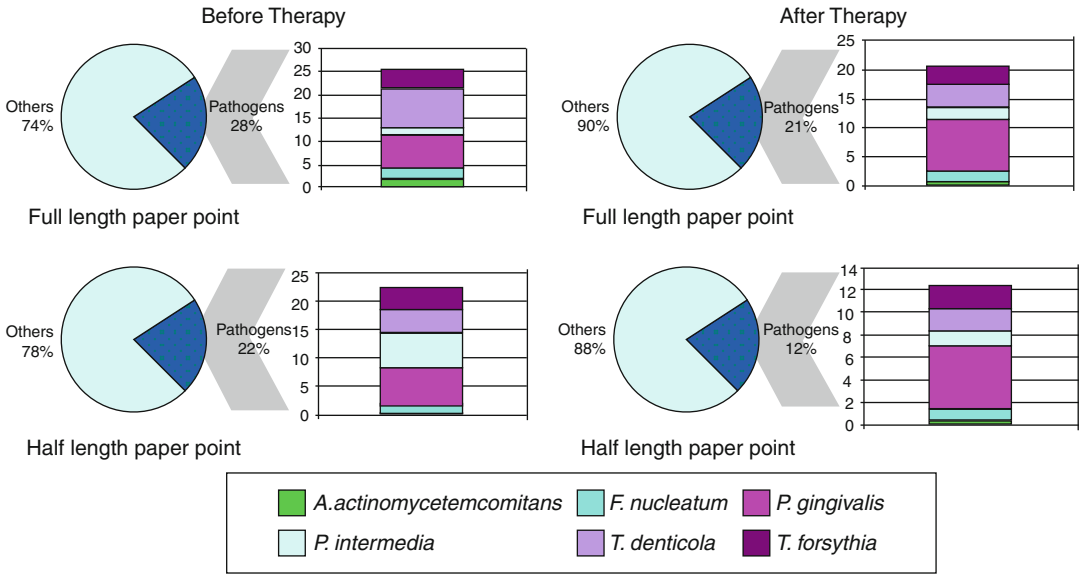


Fig. 12.28 Means of proportions of full length paper point samples and half length paper point samples in group B (Jervøe-Storm et al. 2010. Reprinted with permission from Springer Science+ Business Media)

two sampling techniques in group A, the plaque composition as to target pathogens was fairly similar. In group B, full-length paper points sampled always higher TTPP than half-length paper points at any time point (before/after therapy: full length 26%/21%, half length 22%/12%; **Fig. 12.28**). A more pronounced reduction of TTPP was recorded with half-length paper points. Also, the plaque composition regarding total target pathogens was similar for both samples. Both for total bacterial count and for single target bacteria, a strong positive correlation was found between paper point half length and full length (**Table 12.2**). The results of this study provide new information on the currently widely accepted paper point sampling technique in microbiological diagnostic assays of periodontal lesions. Even though paper points inserted to full length of the probing depth aiming at the apical region collected higher amounts of total bacteria than paper points inserted to the coronal half only, probably due to their greater sampling volume, there was a relatively good agreement between both samples for the recovery of selected target pathogens. However, after periodontal therapy, there might be an indication

that paper points inserted to the full depth of the periodontal pocket would be more suitable to collect periodontopathogenic bacteria, probably due to the reduced amounts of microorganisms as a result of instrumentation, demanding as much absorbing material in the pocket as possible. These findings seem to indicate that the anaerobic periodontal pathogens do not primarily inhabit deeper regions, but may be present in similar proportions throughout the various depths of the periodontal pocket (Jervøe-Storm et al. 2010).

12.4.1.2 The Number of Sites to Be Sampled and the Selection of These Sites Strategies

When taking subgingival samples, the main aim is to obtain a representative sample of the subgingival microbiota from the patient's mouth, and therefore, the best sampling strategy would be the one that achieves this goal more accurately. The number of sites to be sampled and the selection of these sites are two of the parameters that may influence this goal. Sites with the deepest pocket and bleeding on probing are those with the highest chance of harbouring *P. gingivalis*

Table 12.2 Spearman correlation and Kappa coefficients between half and full length paper point samples in groups A and B at all time points ($n=40$) (Jervøe-Storm et al. 2010) (Reprinted with permission from Springer Science + Business Media)

	Spearman coefficient	P value	Kappa	95% CI		Evaluation
				Lower	Upper	
Total bacteria counts	0.588	0.000				
<i>A. actinomycetemcomitans</i>	0.778	0.000	0.807	0.597	1.000	Excellent agreement
<i>F. nucleatum</i>	0.772	0.000	0.573	0.292	0.854	Fair agreement
<i>P. gingivalis</i>	0.824	0.000	0.733	0.488	0.979	Good agreement
<i>P. intermedia</i>	0.756	0.000	0.480	0.205	0.756	Fair agreement
<i>T. denticola</i>	0.814	0.000	0.807	0.597	1.000	Excellent agreement
<i>T. forsythia</i>	0.695	0.000	0.692	0.403	0.972	Good agreement

and *A. actinomycetemcomitans*. In sites with bleeding, the chances of detecting *P. gingivalis* are 4.3 times higher than in non-bleeding sites (Mombelli et al. 1991a, b, 1994a, b). These differences in the composition of the subgingival microbiota have been related, in part, to local environmental factors such as Red-Ox potential and anaerobic conditions, and the availability of nutrients originated from blood products such as growth factors. Based on these data, the most common sampling strategy consists in selecting the sites with the deepest probing depth and bleeding in each quadrant and processing the obtained material as a pooled sample (Casas et al. 2007 and references therein).

Given the similarities in the clinical evidence of periodontitis, it was shown that the presence and levels of 37 species commonly studied in periodontitis are similar, with **no differences regarding sample location (mesio-buccal through disto-lingual/palatal) or between molar, premolar, and incisor/cuspid subgingival sites** (Persson et al. 2008). However, in some studies, 1–4 plaque samples have been taken from each subject (Van der Weijden et al. 1994; Conrads et al. 1996; Meurman et al. 1997), whereas other studies sampled each tooth in the oral cavity (Christersson et al. 1992; Preus et al. 1994; Nakagawa et al. 1996; Dibart et al. 1998). The collection of subgingival plaque samples from many sites using these instruments is time consuming and may also fail to detect some important organisms when limited gingival sites are available (Casas et al. 2007).

Recently, Casas et al. (2007) tested whether different sampling strategies would influence the microbiologic outcomes by assessing the bacterial load and composition of the subgingival microbiota by means of anaerobic culturing before and after periodontal treatment. The results from this study showed that different sampling strategies may render significantly different microbiologic results from the same patient's mouth. Before treatment, although the results in the frequency of detection of putative periodontal pathogens or their proportion in the microbiota was similar irrespective of the sampling strategy, the presence of a relevant number of false negatives for strategies with one or two sample sites versus four sites should be considered when selecting a sampling strategy. In the evaluation of microbiologic outcomes after periodontal therapy, sampling four sites was more accurate than sampling two sites or one site. The selection criteria and the number of sites selected for sampling the subgingival biofilm for anaerobic culturing in patients with generalized chronic periodontitis may influence the microbiologic outcomes, especially after treatment.

12.4.1.3 The Reproducibility of Subgingival Plaque-Sampling Methods

Although multiple studies in the literature have compared different sampling techniques and different sampling strategies, few studies have evaluated the reproducibility of subgingival

plaque-sampling methods (Dahlen et al. 1990; Mombelli et al. 1989; Mousques et al. 1980; Teles et al. 2008).

Dahlen et al. (1990) performed duplicate microbiological samples taken 1 week apart using the paper point technique from a total of 112 untreated periodontal pockets greater than 6 mm deep in 16 adult periodontal patients. Duplicate samples were also obtained from these sites 6 months following a therapy of oral hygiene instruction and supra- and subgingival debridement. The reproducibility of the total viable counts and the reproducibility of the proportions of various groups or species of microorganisms were studied from these duplicate samples. Using cultural techniques for bacterial identification, they found 'acceptable' levels of reproducibility for *Aggregatibacter actinomycetemcomitans* (previously *Actinobacillus actinomycetemcomitans*) and *P. gingivalis*. However, 'unacceptable' levels of reproducibility were observed for total viable counts and for the other investigated bacterial taxa such as *P. intermedia*.

Mombelli et al. (1989) examined the reproducibility of sampling in a total of 109 sites that were sampled repeatedly with sterile paper points at an interval of 7–10 days in 24 patients suffering from periodontal disease and 12 edentulous patients wearing successful and failing osseointegrated titanium implants. The samples were enumerated using cultural techniques and by darkfield microscopy. The standard deviation of proportional differences between first and second samples ranged between 6.4% (fusiform organisms) and 17.2% (coccoid cells) for darkfield parameters, between 4.3% (*P. melaninogenicus*) and 14.0% (*P. gingivalis*) for selected bacterial species, and between 6.9% (gram-negative anaerobic cocci) and 24.0% (gram-positive facultative cocci) for bacterial groups classified according to gram stain characteristics and atmospheric growth conditions.

Teles et al. (2008) examined the reproducibility of curette sampling of subgingival biofilms. Seven subgingival biofilm samples were taken successively, using a curette, from each of 80

sites and individually analysed for their content of 40 bacterial species using checkerboard DNA–DNA hybridization. One healthy site was sampled in each of 20 periodontally healthy subjects, and one sulcus/pocket of ≤ 3 , 4–5, and ≥ 6 mm was sampled in each of 20 subjects with chronic periodontitis. It was found that one curette stroke was not enough to remove all of the plaque present in one site. One could obtain visually detectable plaque samples up to the fourth stroke in most of the moderate and deep periodontal pockets. It is conceivable that one cannot empty the periodontal pocket because removal of plaque might create an influx of plaque from adjacent areas. It was also demonstrated that each of the seven strokes from the same site provided samples that were very similar in terms of the proportions of the test species, although the bacterial counts decreased. Further, samples were much more similar within the same site than samples taken from different sites, as demonstrated by the coefficient of variation and the similarity coefficients. The median coefficient of variation for individual species in the same site was 0.79 (95% confidence interval [CI]: 0.76–0.82) compared to 1.76 (95% CI: 1.69–1.82) in samples from different sites. The within-site mean minimum similarity coefficient ($\pm$ SEM) was $51.2\% \pm 2.2\%$, and it was $27.9\% \pm 0.3\%$ between sites. Thus, the findings of the current study suggested that the use of curettes provided a reliable and reproducible method to obtain subgingival samples.

There is ample evidence that the periodontal pocket represents the location in which disease activity primarily takes place and that microorganisms in the periodontal pocket are the paramount etiologic factor of periodontal disease. Therefore, monitoring the subgingival region for the occurrence of periodontal pathogens has been considered reasonable for a long time, and most of the studies evaluating the prevalence of periodontal pathogens or microbiological effects of periodontal therapy have used a subgingival sampling strategy to describe a microbiological colonization. However, it has been demonstrated that periodontal pathogens

can also be detected in other intraoral habitats, for example, the supragingival region, mucous membranes, or saliva. Moreover, periodontal pathogens seem to be easily transferred from one intraoral niche to another, thus indicating that the whole intraoral cavity constitutes the habitat for periodontal pathogens (Beikler et al. 2006 and references therein). Consequently, microbiological sampling of the subgingival region alone may not completely reflect the intraoral colonization and, therefore, may underestimate the detection frequency of a specific pathogen. Umeda et al. (1998) compared the presence of six periodontopathic bacteria in whole saliva and subgingival plaque of 202 subjects. The test bacteria were identified using a 16S rRNA-based PCR detection method. Each study subject contributed a whole saliva sample and a paper point sample pooled from the deepest periodontal pocket in each quadrant of the dentition. The results indicate that whole saliva is superior to pooled periodontal pocket samples to detect *P. gingivalis*, *P. intermedia*, *P. nigrescens*, and *T. denticola* in the oral cavity. The detection of oral *A. actinomycetemcomitans* and *B. forsythus* with reasonably good accuracy may require both whole saliva and periodontal pocket samples. With the exception of *T. denticola*, this has been confirmed by Beikler et al. (2006) who found that analysing only a single sample of each habitat does not result in a sufficient probability for the intraoral detection of assessed pathogens. It was suggested that for determining the intraoral carrier state of patients with periodontitis, a combined sample of supra- and subgingival plaque taken from the deepest periodontal pocket in each sextant may yield the most reliable result.

The use of saliva for diagnostic purposes has been the subject of considerable research (Mandel 1990). Saliva is easy to obtain and contains bacteria from different oral sites including oral mucosal sites and supra- and subgingival plaque. Consequently, studies have evaluated the presence and levels of bacteria in saliva in relation to the periodontal status (Boutaga et al. 2007). Asikainen et al. (1991) used the culture technique

to investigate the presence of *A. actinomycetemcomitans* in stimulated and unstimulated saliva samples and found these pathogens in 69.9% and 35.9% of the samples, respectively, when *A. actinomycetemcomitans* was detected subgingivally. Umeda et al. (1998) used the polymerase chain reaction (PCR) technique to detect periodontal pathogens in pooled subgingival plaque and whole saliva samples. They found that whole saliva samples were more frequently positive for *P. gingivalis* and *Treponema denticola* than subgingival plaque samples, whereas saliva samples underestimated the presence of *A. actinomycetemcomitans* and *T. forsythensis* compared to subgingival samples. Boutaga et al. (2007) demonstrated that rapid detection and quantification of periodontal pathogens in mouthwash samples are possible by real-time PCR. The procedure is significantly less time consuming than subgingival sampling with paper points. It was suggested that this approach to detect major periodontal pathogens in mouthwash samples may simplify microbial diagnosis in periodontitis patients and may be used to monitor periodontal treatment (Boutaga et al. 2007).

12.4.2 Microscopic Methods

Darkfield and phase contrast microscopy assays bacterial size, shape, and motility. These assays are very inexpensive and take only a few minutes to perform in the office. They can differentiate plaque samples as health-associated plaque characterized by small numbers of mainly non-motile cocci and as disease-associated plaque characterized by a large number of bacterial morphotypes including small, medium, and large spirochetes and motile and curved rods (Listgarten and Levin 1981; Zambon 1997).

These visual techniques cannot identify individual species but can be useful in observing a shift in the composition of the plaque because the infected site is densely populated by motile and rodlike organisms, whereas the healthy one is sparsely populated by coccal and non-motile bacteria (D'Ercole et al. 2008).

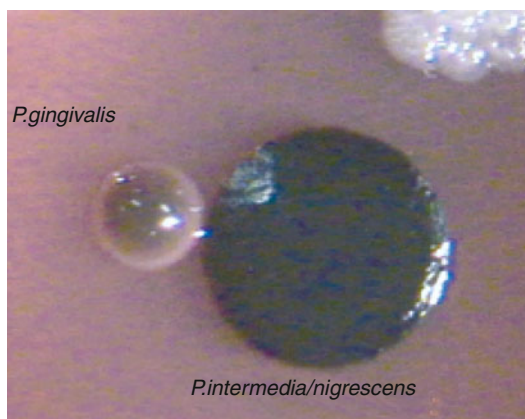


Fig. 12.29 Colony of *Porphyromonas gingivalis*, close to another colony of *Prevotella intermedia/nigrescens*, growing on blood agar medium, supplemented with haemin and menadion (Sanz et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

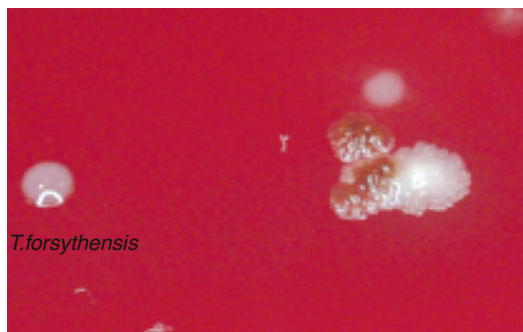


Fig. 12.30 Colony of *Tannerella forsythensis*, growing on blood agar medium (Sanz et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

12.4.3 Culture Methods

Bacterial culture is the ‘gold standard’ microbiological assay against which other tests are compared and validated. These tests are expensive, are performed in a reference laboratory, and can give information which other tests do not provide—antibiotic susceptibility and resistance. This information is valuable in treatment planning and in choosing an appropriate antibiotic (**Fig. 12.29**) (Zambon 1997).

Generally, subgingival plaque samples are cultivated anaerobically, and by using selective and non-selective media, together with several biochemical and physical tests, the different putative pathogens can be identified. However, culture techniques have significant shortcomings. Culture methods can only grow viable bacteria; therefore, strict sampling and transport conditions are essential. Moreover, some of the putative pathogens, such as *Treponema* sp. and *Tannerella forsythensis*, are very fastidious and difficult to culture (**Fig. 12.30**). The sensitivity of bacterial culturing can be rather low, especially for non-selective media, with detection limits averaging 10^3 – 10^4 bacterial cells, and therefore, low numbers of a specific pathogen in

a subgingival sample will be undetected. However, the most important drawback is that culture requires specific laboratory equipment and experienced personnel, besides being relatively time consuming and expensive (Sanz et al. 2004).

12.4.4 Enzymatic Methods

An enzymatic assay has been developed that detects bacteria that possess trypsin-like enzymes such as *T. forsythensis*, *Treponema denticola*, and *P. gingivalis*. When a plaque sample containing any combination of these three bacteria is placed on a paper strip impregnated with a colourless substrate *N*-benzoyl-DL-arginine-2-naphthylamide (BANA), the BANA substrate breaks down and produces a blue-black colour whose intensity is proportional to the total amount of the three organisms (Perioscan, Oral-B Laboratories, Redwood City, CA 94065, USA) (**Fig. 12.31**) (Loomer 2004). While this chairside test was unable to distinguish between the relative proportions of the three bacteria and cannot identify the presence of other oral microorganisms, it showed a statistically significant correlation between increasing probing depth and a positive BANA test. BANA test detected the presence of BANA-positive microorganisms at sites of <3-mm probing depth in a statistically



Fig. 12.31 The BANA kit

significant proportion (Grisi et al. 1998). The BANA test revealed a sensitivity of 99% and a specificity of 55% when related to the clinical diagnosis. The probability that the tests agreed with the clinical outcome after treatment was calculated as 52% (Hemmings et al. 1997). It may have more value when performed in combination with other chairside microbiological tests, such as microscopy (Loomer 2004).

12.4.5 Immunological Methods

Immunoassays use monoclonal or polyclonal antibodies against species-specific antigens, and the identification of these reactions allows the detection of target bacteria. This reaction can be visualized using a variety of techniques and reactions, as direct (DFA) and indirect (IFA) immunofluorescent microscopy assays, flow cytometry, enzyme-linked immunosorbent assay (ELISA), membrane assays, and latex agglutination (D'Ercole et al. 2008).

In periodontology, the specific antibody dosage in serum and saliva has also been used. In this case, a radioimmunoassay (RIA) has been used, in which the antibody is labelled with a radioisotope ^{125}I (D'Ercole et al. 2008).

Comparative analyses of subgingival plaque samples with culture and indirect immunofluorescence (Fig. 12.32) clearly document the

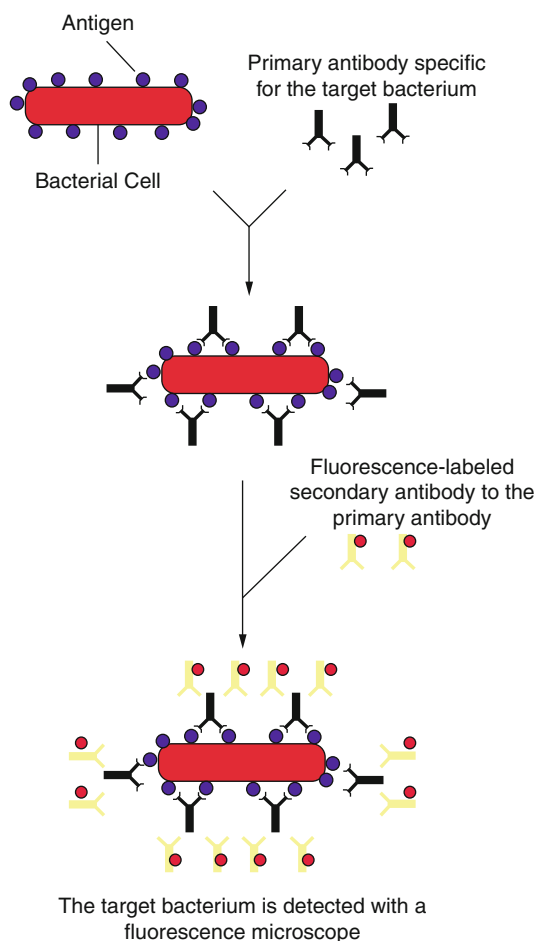


Fig. 12.32 Indirect immunofluorescence detection (Chen and Slots 1999. Reprinted with permission from John Wiley & Sons, Inc.)

usefulness of the fluorescence approach. Bonta et al. (1985) detected the prevalence of *A. actinomycetemcomitans* using immunofluorescence and culture techniques. Utilizing polyclonal antiserum, the sensitivity of indirect immunofluorescence microscopy was 100% in comparison with that of bacterial culture on non-selective medium and 96% as compared with that of bacterial culture on selective medium. The sensitivity of the BA-28 monoclonal antibody was found 90% in comparison with that of non-selective medium and 82% as compared with that of selective medium. Immunofluorescence using either polyclonal antisera or monoclonal antibodies exhibited high levels of specificity—8% as compared with that of culture on non-selective media and 92% with selective media (Bonta et al. 1985). Kamma et al. (2004) reported a 75.9% agreement between the two techniques for *A. actinomycetemcomitans*.

Agreement for the detection of *P. gingivalis* using immunofluorescence and culture techniques was very high at 88.1% in the study performed by Kamma et al. (2004). Similar results were reported by Slots et al. (1985) who found an 80.5% agreement between culture and immunofluorescence for *P. gingivalis* detection. The sensitivity of immunofluorescence for detecting *P. gingivalis* was 91% and the specificity 80%. In contrast, Zambon et al. (1985) found *P. gingivalis* twice as frequently with immunofluorescence than with culture. Gmür (1988) reported that culture consistently yielded lower numbers of *P. gingivalis* cells than did immunofluorescence. For *P. intermedia*/*P. nigrescens*, the agreement was also high at 86.5% (Kamma et al. 2004), which was in accordance with the findings of Gmür (1988) who detected a remarkably similar number of cells by both techniques. The agreement between immunofluorescence and culture techniques for *T. forsythia* in a study performed by Kamma et al. (2004) was 85.2%, whereas Gmür (1988) found a greater discrepancy because culture failed to isolate it.

Cytofluorography or flow cytometry for the rapid identification of oral bacteria involves labelling bacterial cells from a patient plaque sample with both species-specific antibodies and

a second fluorescein-conjugated antibody. The suspension is then introduced into the flow cytometer, which separates the bacterial cells into an almost single-cell suspension by means of a laminar flow through a narrow tube (Kamiya et al. 1994). The sophistication and cost involved in this procedure preclude its wide usage (Sanz et al. 2004).

ELISA is similar in principle to other radioimmunoassays, but instead of the radioisotope, an enzymatically derived colour reaction is substituted as the label. The intensity of the colour depends on the concentration of the antigen, and it is usually read photometrically for optimal quantification. ELISA has been used primarily to detect serum antibodies to periodontal pathogens; however, it has also been used in research studies to quantify specific pathogens in subgingival samples using specific monoclonal antibodies. A membrane immunoassay has been adapted for chairside clinical diagnostic use and has been marketed (Evalusite®, Eastman Kodak, Rochester, NY, USA). It involves linkage between the antigen and a membrane-bound antibody to form an immunocomplex that is later revealed through a colorimetric reaction (Sanz et al. 2004). It demonstrated a high specificity of the Evalusite® Test for periodontally diseased sites and indicated that, in terms of clinical significance, the sensitivity threshold for the Evalusite® Test corresponds to the presence of $>10^4$ cultivable cells of *A. actinomycetemcomitans*, *P. gingivalis*, and *P. intermedia* (Boyer et al. 1996; Riep et al. 1999; Chaves et al. 2000).

These methods have several important advantages including relatively low cost, rapidity, and no requirement of stringent sampling and transport methodology to assure bacterial viability. However, they require the use of monoclonal antibodies to assure high specificity, and the detection limits are significantly high (10^3 – 10^4 bacterial cells). Moreover, they are limited to the number of antibodies tested, lack the validity of well-controlled clinical studies, and above all do not permit evaluation of antibiotic susceptibility of the flora (D'Ercole et al. 2008).

12.4.6 Nucleic Acid Probe Method

Nucleic acid probes consist of nucleic acid sequences that are labelled with a radioactive or enzymatic–colorimetric marker that bind to complementary nucleic acid sequences on corresponding microorganisms. The probe sequences may be whole **genomic, randomly cloned sequences of nucleic acids, or synthetic oligonucleotides** (also known as 16S rRNA probes). Of the three, it is the oligonucleotide probes that have the greatest specificity and lowest cross-reactivity because they target genes specific to a bacterial species. Whole genomic probes and random-cloned probes may contain sequences common to multiple species (Loomer 2004).

When DNA probe analyses were compared to cultural methods for microbiological analysis of plaque, the DNA probe analysis demonstrated 100% effectiveness in detecting *A. actinomycetemcomitans* and *B. intermedius* and 91% effectiveness in detecting *B. gingivalis* at culture-positive levels (greater than or equal to 10^3 cells). In addition, probe assays frequently identified these pathogens in samples that were culture negative. Probe analysis revealed a better correlation between presence of a pathogen and clinical evidence of disease on an individual patient basis. In addition, viability of microorganisms is not a requirement for nucleic acid probe analysis that may be an advantage when plaque sample analysis may be delayed because of lengthy transportation from clinic to laboratory (Savitt et al. 1988). The DNA probe method used in the identification of *Tannerella forsythensis* and *Porphyromonas gingivalis* demonstrated a detection threshold of 10^5 cells, relatively high when compared to the culture method (Tsai et al. 2003).

12.4.7 DNA–DNA Hybridization Methods

12.4.7.1 Fluorescence In Situ Hybridization

The fluorescence in situ hybridization (FISH), or more specifically whole-cell hybridization tech-

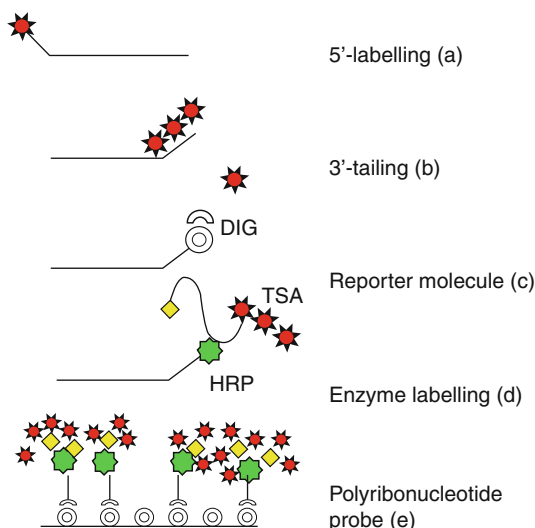


Fig. 12.33 Direct (a, b) and indirect (c–e) labelling of probes using a reporter molecule like digoxigenin (DIG) that is then detected by fluorescent antibody, horseradish peroxidase (HRP) that used fluorescein–tyramide as substrate for the enzymatic signal amplification [tyramide signal amplification (TSA)], or with the combined use of polyribonucleotide probes, internally labelled with digoxigenin, with the tyramide signal amplification systems (Bottari et al. 2006. Reprinted with permission from Springer Science + Business Media)

nique, detects nucleic acid sequences by a fluorescently labelled probe that hybridizes specifically to its complementary target sequence within the intact cell. The procedure includes the following steps: (1) fixation of the specimen; (2) preparation of the sample, possibly including specific pretreatment steps; (3) hybridization with the respective probes for detecting the respective target sequences; (4) washing steps to remove unbound probes; and (5) mounting, visualization, and documentation of results. In the following sections, we concentrate on the most commonly used techniques only (Gmür and Thurnheer 2002). In the early 1990s, improved labelling of synthetic, single-stranded DNA probes allowed the chemical preparation of hybridization probes carrying enough fluorescent molecules to allow direct detection. Nowadays, there are different ways of labelling (Fig. 12.33) (Bottari et al. 2006). A large number of intact ribosomes representing the biological activity of

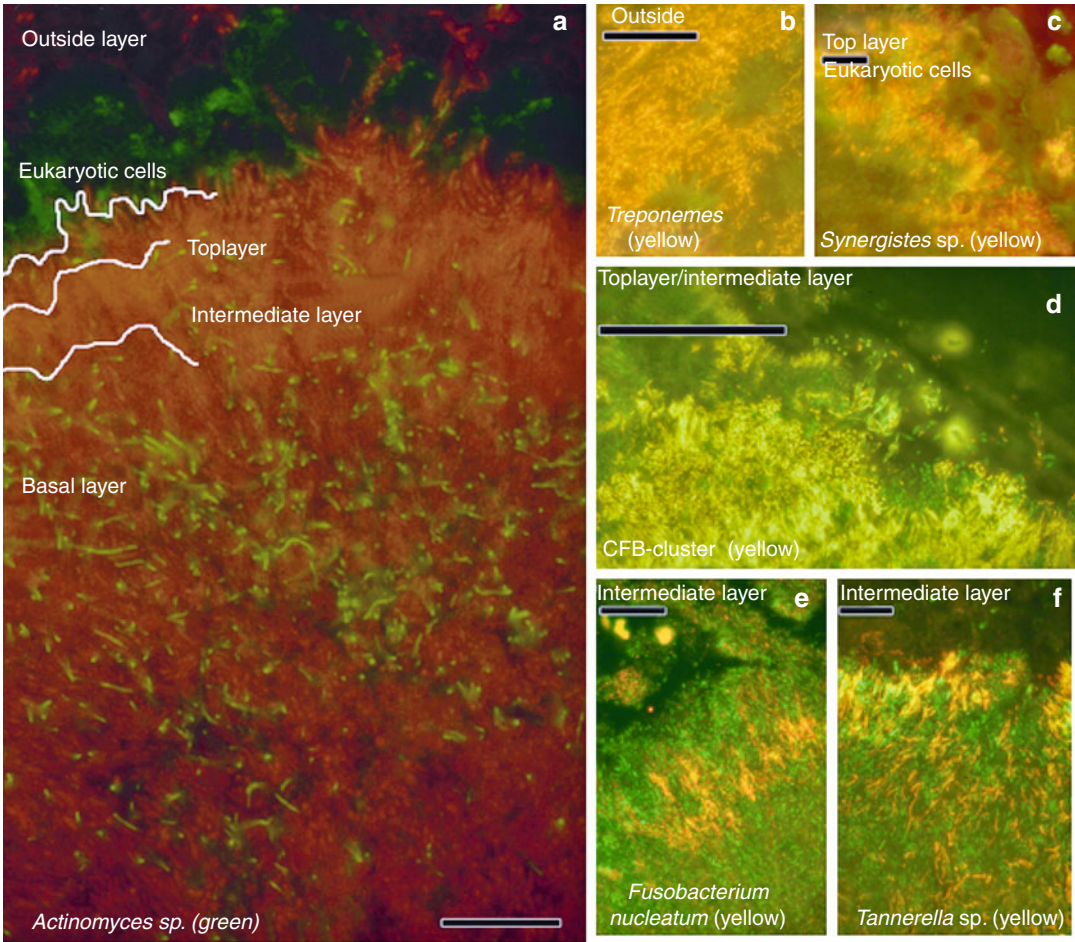


Fig. 12.34 Localization of the most abundant species in subgingival biofilms. (a) Overview of the subgingival biofilm with *Actinomyces* sp. (green bacteria), bacteria (red), and eukaryotic cells (large green cells on top). (b) *Spirochaetes* (yellow) outside the biofilm. (c) Detail of *Synergistetes* (yellow) in the top layer in close proximity to eukaryotic cells (green). (d) CFB-cluster (yellow) in

the top and intermediate layer. (e) *F. nucleatum* in the intermediate layer. (f) *Tannerella* sp. (yellow) in the intermediate layer. Each panel is double stained with probe EUB338 labelled with FITC or Cy3. The yellow colour results from the simultaneous staining with FITC- and Cy3-labelled probes. Bars are 10 µm (doi:10.1371/journal.pone.0009321.g002; Zijng et al. 2010)

the tested cells are a prerequisite for this method, so that apparently only vital bacteria are stained (Amann et al. 1990, 1995; Amann 1995; Hannig et al. 2007, 2010).

FISH has been successfully used in combination with confocal laser scanning microscopy (CLSM) and epifluorescence microscopy for the visualization to determine the spatial configuration and demonstrate the morphology of individual bacterial cells in complex natural

communities, such as dental plaque (Figs. 12.34 and 12.35) (Thurnheer et al. 2001; Al-Ahmad et al. 2007, 2009, 2010; Dige et al. 2007, 2009; Hannig et al. 2007; Colombo et al. 2007; Zijng et al. 2010).

12.4.7.2 Checkerboard Hybridization

In order to make definitive bacterial associations with oral health and disease states, the microbial profiles of large numbers of clinical samples must

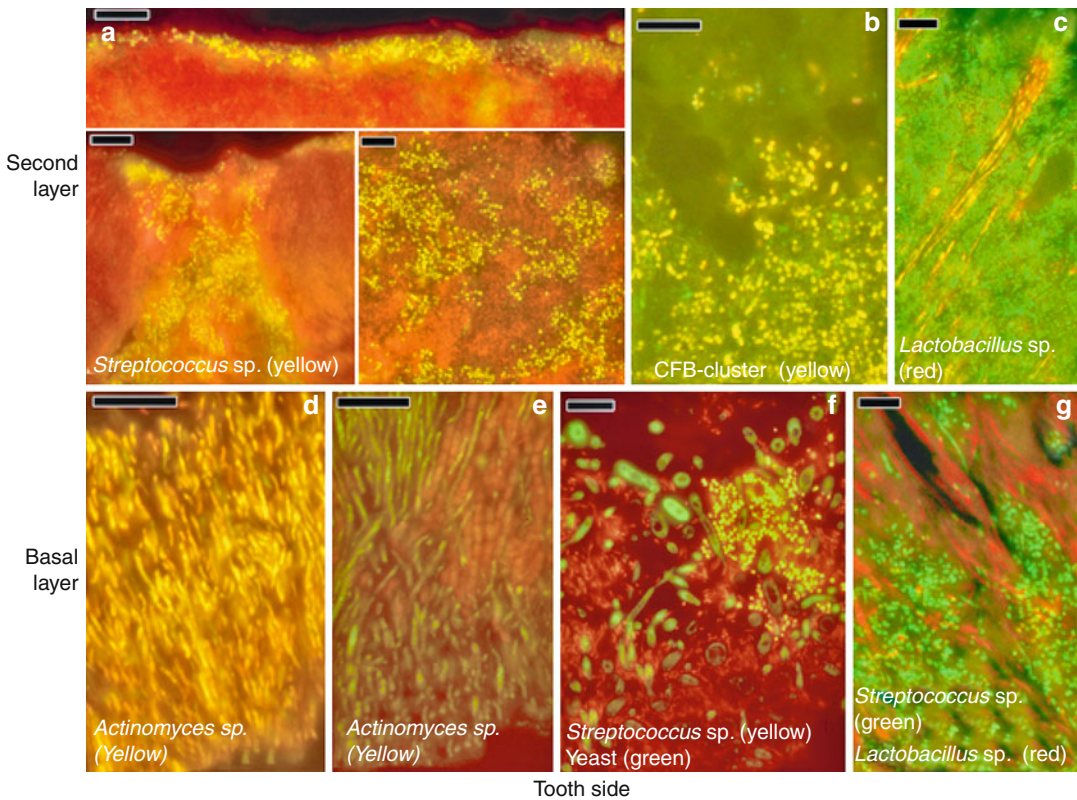


Fig. 12.35 Localization of the most abundant species in supragingival biofilms. *Streptococcus* sp. (yellow) form a thin band on top of the biofilm (a1), almost engulfing in the biofilm (a2) or present as small cells scattered through the top layer of the biofilm (a3). (b) Cells from the CFB-cluster of bacteria in the top layer of the biofilm, without defined structure. (c) *Lactobacillus* sp. (red) forming long strings through the top layer. (d) *Actinomyces* sp. (yellow) plaque attached to the tooth. (e) *Actinomyces* sp. (green)

and cocci forming initial plaque. (f) Multispecies initial plaque composed of *Streptococcus* sp. (yellow), yeast cells (green), and bacteria unidentified (red). (g) *Streptococcus* sp. (green) and *Lactobacillus* sp. (red) forming initial plaque. Black holes might be channels through the biofilm. Panels a–c, e, f are double stained with probe EUB338 labelled with FITC or Cy3. Bars are 10 μ m (doi:10.1371/journal.pone.0009321.g003; Zijlge et al. 2010)

be determined. There are two types of checkerboard hybridization: one utilizing whole genomic DNA probes that are hybridized to sample DNA on the membrane (Socransky et al. 1994) and the other utilizing labelled 16S ribosomal RNA amplicons that are hybridized to 16S ribosomal RNA-based probes that are on the membrane (Paster et al. 1998). This latter method has been referred to as reverse-capture, 16S ribosomal RNA-based oligonucleotide checkerboard hybridization. In both methods, hybridization signals are typically detected using chemifluorescence procedures (Paster and Dewhirst 2009).

Although there is potential use for checkerboard hybridization as a diagnostic tool, the checkerboard hybridization methodology is used more routinely for checkerboard hybridization has used whole genomic DNA probes to study the roles of bacteria in oral health and disease (Paster and Dewhirst 2009). In a landmark paper, Socransky et al. (1994) introduced this assay that uses whole genomic, digoxigenin-labelled DNA probes and facilitates rapid processing of large numbers of plaque samples with respect to a multiple hybridization for up to 40 oral species in one single test. The DNA probes used in this

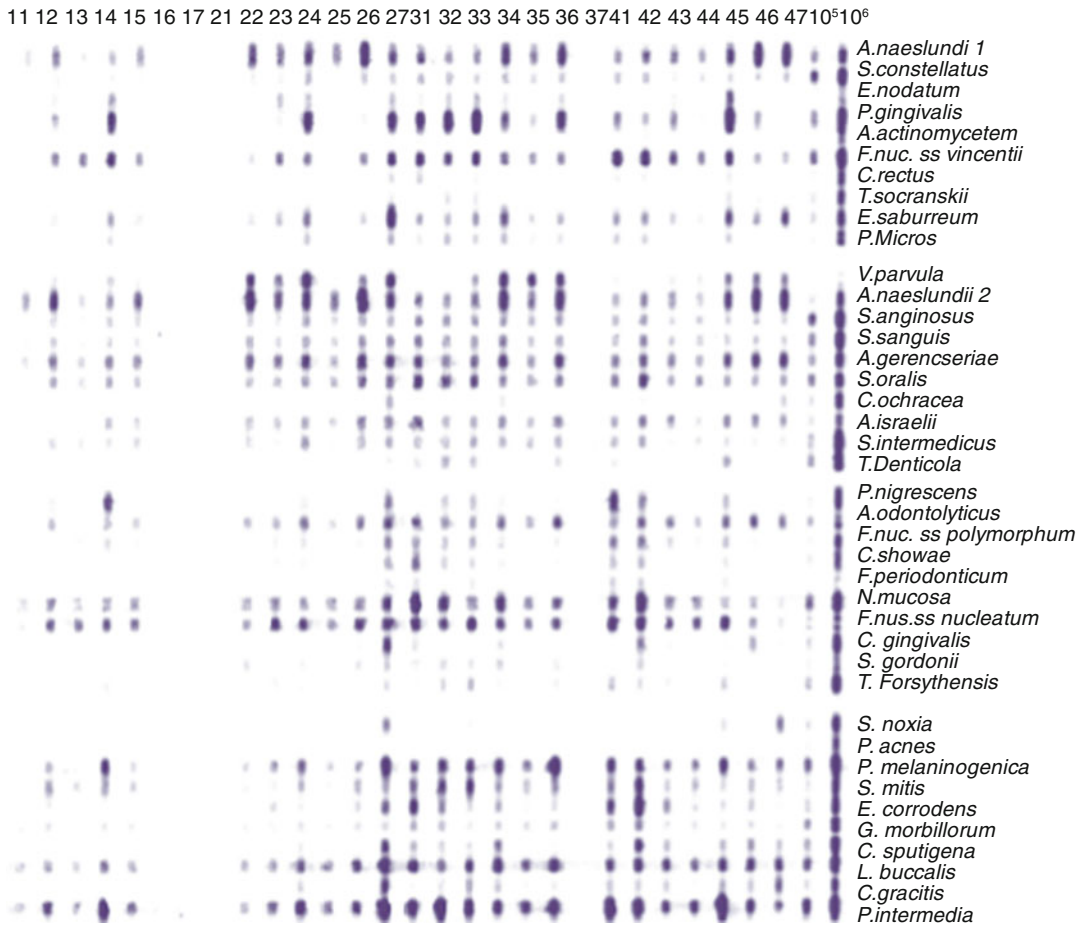


Fig. 12.36 Example of checkerboard DNA–DNA hybridization being used to detect 40 bacterial species in 28 subgingival plaque samples from a single patient. The vertical lanes are the plaque samples numbered from 11 (right maxillary central incisor) to 47 (right mandibular second molar). In this subject, teeth 16, 17, 21, and 37 were missing. The two vertical lanes on the right are standards containing either 10^5 or 10^6 cells of each test species. The horizontal lanes contained the indicated DNA probes in hybridization buffer. A signal at the intersection of the vertical and horizontal lanes indicates the presence of a species. The intensity of the signal is related to the number of organisms of that species in the sample. In brief, samples of plaque were placed into individual Eppendorf tubes and the DNA released from the microor-

ganisms by boiling in NaOH. After neutralization, the released DNA was transferred to the surface of a nylon membrane using the 30 channels of a Minislot device (Immunetics). The DNA was fixed to the membrane by ultraviolet light and baking and placed in a Miniblotter 45 (Immunetics) with the lanes of DNA at right angles to the 45 channels of the Miniblotter device. Whole genomic DNA probes labeled with digoxigenin were placed in hybridization buffer into 40 of the lanes and hybridized overnight. After stringency washing, the signals were detected using phosphatase-conjugated antibody to digoxigenin and chemifluorescence substrates. Signals were compared to the standards using a Storm Fluorimager and converted to counts (Socransky et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

technology are commonly adjusted to permit detection of 10^4 cells of each species, but can be adjusted to detect 10^3 cells. The method requires sophisticated laboratory equipment and expertise, and it is highly specific. These factors have not led to generalization of this assay for diagnostic pur-

poses. It is particularly applicable, however, for epidemiological research and ecological studies, since it does not require viable bacteria and allows for the assessment of a large number of plaque samples and multitude of species (Fig. 12.36) (Sanz et al. 2004 and reference therein).

Papapanou et al. (1997) compared the 'checkerboard' DNA–DNA hybridization methodology with culture techniques for the analysis of the composition of the subgingival microbiota. Seventy subjects, presenting with a variety of periodontal conditions, contributed a total of 283 subgingival plaque samples analysed with respect to the following species: *Porphyromonas gingivalis*, *Prevotella intermedia/Prevotella nigrescens*, *Fusobacterium nucleatum*, *Campylobacter rectus*, *Eikenella corrodens*, *Bacteroides forsythus*, *Actinobacillus actinomycetemcomitans*, *Streptococcus sanguis*, and *Streptococcus mutans*. It was found that the checkerboard technology resulted in higher prevalence figures for half of the species tested when compared to culture data. If the latter were used as the reference, checkerboard detection sensitivities ranged from 0.17 to 0.86, specificities from 0.17 to 1.0, and diagnostic accuracies from 0.51 to 0.81, depending on bacterial species. The use of the checkerboard data as the reference resulted in detection sensitivities for the culture procedures between 0.24 and 1.0 and specificities between 0.21 and 0.87. The checkerboard methodology resulted in statistically significant higher bacterial counts for the majority of the species. It was further observed that, for most species, the higher the total number colony-forming units in the sample, the higher the discrepancy between the results obtained by the two techniques.

Siqueira et al. (2002) compared 16S rDNA-based PCR and the checkerboard technique for the content of *A. actinomycetemcomitans*, *T. forsythia*, *Parvimonas micra*, *Porphyromonas endodontalis*, *P. gingivalis*, and *T. denticola* in samples from root canals. On the whole, matching results between the two molecular methods ranged from 60% to 97.5%, depending on the target species. The major discrepancies between the methods comprised a number of PCR-positive but checkerboard-negative results.

Recently, Haffajee et al. (2009) compared polymerase chain reaction (PCR) with subsequent reverse hybridization (micro-IDent test) and checkerboard DNA–DNA hybridization for the identification of 13 bacterial species in sub-

gingival plaque samples. Three hundred and fifty samples were evaluated using both techniques. Regression analysis indicated that 10/13 test species showed significant positive correlations between the counts determined by checkerboard analysis and levels determined by the PCR-based test after adjusting for 13 comparisons. The highest rank correlations of 0.58, 0.49, and 0.46 were seen for *Treponema denticola*, *Fusobacterium nucleatum*, and *Eubacterium nodatum*, respectively ($P < 0.0001$). Both tests could distinguish samples from healthy and periodontitis subjects.

The reverse-capture checkerboard hybridization method has also been used to determine the role of bacteria in cariology (Lima et al. 2011), endodontics (Rôças and Siqueira 2008, 2010; Siqueira et al. 2009; Rôças et al. 2010), and periodontology (Paster et al. 2002; Aas et al. 2007) research.

12.4.7.3 Oligonucleotide Microarray Technology

As an extension of the 16S ribosomal RNA-based, reverse-capture DNA–DNA checkerboard hybridization, the human oral microbe identification microarray was developed in order to examine the complex oral microbial diversity in a single hybridization reaction on glass slides (Paster and Dewhirst 2009), termed the Human Oral Microbe Identification Microarray (HOMIM). Briefly, 16S rRNA-based oligonucleotide probes are covalently attached to an aldehyde-coated slide through imide links. 16S rRNA genes are polymerase chain reaction amplified from DNA isolated from bacterial cultures or clinical samples (Fig. 12.37). Amplified products are labelled with a fluorescent dye during the polymerase chain reaction step. When labelled bacterial DNA hybridizes to a species-specific spot on the slide, it provides a fluorescent signal that can be read with a microarray laser scanner. At present, the HOMIM allows the identification of 200 predominant cultivable and uncultivable species (Paster et al. 2006). HOMIM was used for analysis of the oral microflora in the elderly (Preza et al. 2008, 2009a, b)

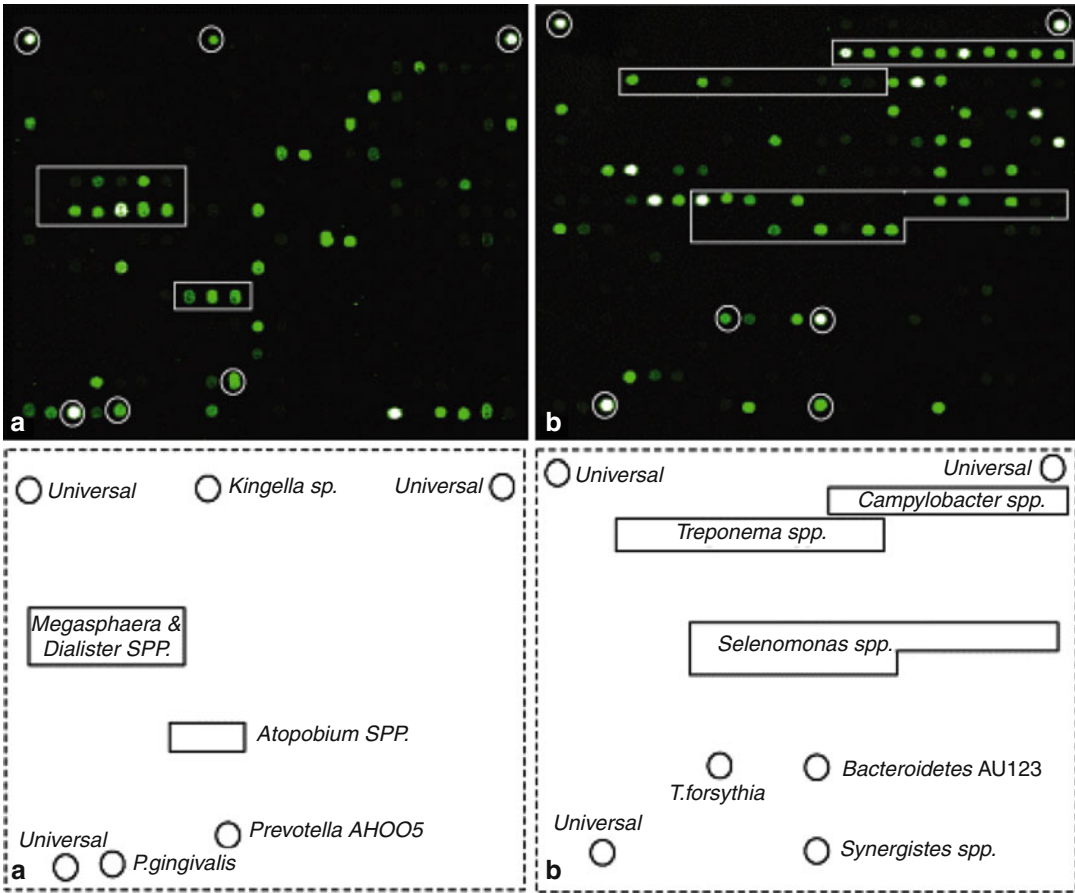


Fig. 12.37 Microbial profiles of subgingival plaque visualized on Human Oral Microbe Identification Microarray (HOMIM); 301 spots (oligonucleotide probes) targeting 200 bacterial species are on each microarray. Probes are organized in phylogenetic groups. Selected species or phylogenetic groups are identified in the figures. Three universal probes are present on each array (shown

above) to ensure proper orientation. Positive hybridization is indicated by a green spot. (a) HOMIM of sample from a healthy site from a periodontally healthy subject; (b) HOMIM of sample from a diseased site from a patient with periodontitis. Differences in patterns are clearly evident (Paster et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

and to compare subgingival microbial profiles of refractory periodontitis, severe periodontitis, and periodontal health (Colombo et al. 2009).

A commercially available microarray system (ParoCheck®; Greiner Bio-One GmbH, Frickenhausen, Germany), which allows the simultaneous detection of up to 20 different oral bacterial species based on species-specific highly conserved regions from the 16S rRNA gene, was used for the detection of the normal microflora of gingival biopsies. The ParoCheck® chip is a coated glass slide with a total of 86 DNA measuring points, which can be evaluated by all com-

mercially available microarray scanners (Eberhard et al. 2008).

Recently, high-density 16S ribosomal RNA-based microarrays have also been developed. The **Phylochip**, developed by the Affymetrix Corporation® (Santa Clara, CA) and Lawrence Berkeley Laboratories (Berkeley, CA) can detect up to 32,000 16S ribosomal RNA phylotypes (Paster and Dewhirst 2009). Huyghe et al. (2008) reported on a similar microarray design, used to study complex bacterial communities, which targets 9,500 16S ribosomal RNA phylotypes (Paster and Dewhirst 2009).

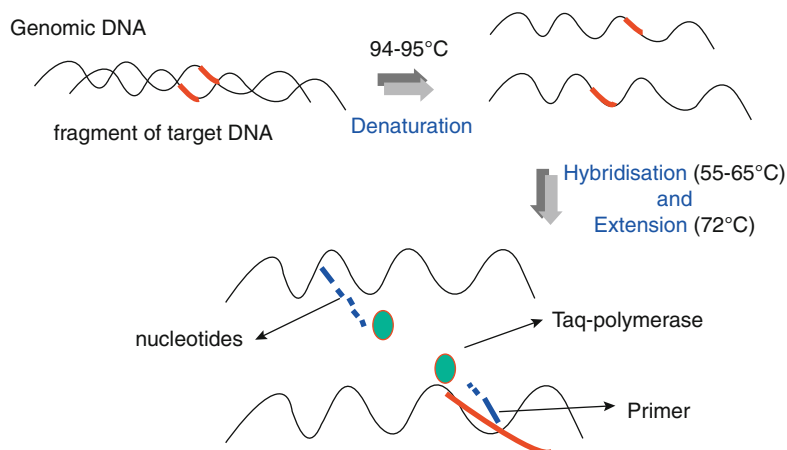


Fig. 12.38 Diagram depicting the polymerase chain reaction process defined by the three stages that comprise this technique: amplification is carried out by a DNA polymerase, once in each cycle. Each cycle includes the denaturalisation or separation of complementary chains,

hybridisation of primers with the original chains and the extension of the primer by the polymerase. Between 30 and 40 cycles are necessary to obtain a significant amount of the studied sequence (Sanz et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

12.4.8 PCR Method

PCR has emerged as the most powerful tool for the amplification of genes and their RNA transcripts. This technique, developed in 1985, is the single technique used almost universally to study DNA and RNA obtained from a variety of tissue sources. PCR allows for obtaining high quantities of DNA in a simplified and automated fashion. PCR typically begins with the isolation of DNA from a fresh tissue specimen. By heating, the complementary double strands of DNA split into single-stranded forms intended to act as the template dictating the nucleotide sequence *in vitro*. Then, the amplification is followed using a DNA polymerase that requires a primer, or known short oligonucleotide sequence corresponding to the border of the region that is amplified. For obtaining amplified fragments of constant length and in high quantities, a second primer, complementary to the opposed chain, must be used to anneal (bind) the template and flank the region of interest. This amplification can be performed several times, known as cycles. In each cycle, the processes of complementary chains denaturation, primer hybridization, and primer extension by means of the polymerase take place (**Fig. 12.38**). With each cycle, there is an exponential increase

in the quantity of DNA. During all these processes, the temperature during the cycle is critical in order to control the double chain denaturation and the stability of the hybridization between the model fragment and the primer (Sanz et al. 2004).

Since first mooted, PCR has evolved from a labour- and time-intensive qualitative technique that relied on the visual interpretation of stained gels to detect the presence of amplification products into today's simple, rapid, and quantitative qPCR, which uses precision optics and DNA-binding fluorescent dyes or fluorescent labels to monitor amplification in real time. This progression was accompanied by prodigious advances in our understanding of the underlying technology as well as the biology it describes (Bustin 2010a).

12.4.8.1 Single Target PCR Applications

Many studies have utilized PCR-based methods to detect specific species directly from oral clinical samples. Species-specific or phylotype-specific PCR primers were designed and subsequently used in highly stringent, individual PCR reactions to establish the prevalence of target species in plaque samples of healthy subjects and of those with periodontal disease. The inves-

tigators confirmed that several additional species, including some that have not yet been cultured *in vitro*, were associated with oral health or periodontitis (Paster and Dewhirst 2009).

Only a small number of PCR assays have used oligonucleotides derived from single copy genes as primers, such as the *P. gingivalis* collagenase prtC gene or the fimbriin fimA gene and the *A. actinomycetemcomitans* leukotoxin lktA gene or the Tf protease prtH gene. 16S rRNA genes as primers, although very specific (cross-reactivity is rare), may not be appropriate for quantitative analysis since there are a variable number of these molecules per cell, which prevents a reproducible quantification. These PCR tests provide only qualitative information (prevalence), and therefore, their use for diagnostic and prognostic purposes in clinical use is limited. Most of these diagnostic tests have been used in ecological studies assessing the prevalence of the target microorganisms in different populations, both in health and disease. Results from these studies are very heterogeneous, showing different prevalences (ranging from 2% to 95% for *A. actinomycetemcomitans*) of the target bacteria in different populations both in health and disease. This heterogeneity might indicate that the high sensitivity of this technique is able to detect low numbers of bacteria that might be irrelevant in terms of pathogenicity or the possibility of different serotypes from the same species being highly prevalent in some populations without leading to pathogenicity. The importance of a quantitative assessment of the target bacteria has led to the recent development of quantitative PCR methods (Sanz et al. 2004).

12.4.8.2 Multiplex PCR

This technique is an expansion of single target PCR methodology in which more than one pair of species-specific primers is used in a single PCR assay and that permits multiple species to be detected simultaneously (Paster and Dewhirst 2009). Such assays have been used to detect several periodontal pathogens, such as *A. actinomycetemcomitans*, *T. forsythia*, and *P. gingivalis*, as well as the presence of herpesviruses, in particular, Epstein–Barr virus and human cytomegalovirus

at the same time (Estrela et al. 2010; De La Garza-Ramos et al. 2008; Ready et al. 2008; Morikawa et al. 2008; Squeri et al. 2006; Santangelo et al. 2004; Tran and Rudney 1996, 1999; García et al. 1998; Wahlfors et al. 1995).

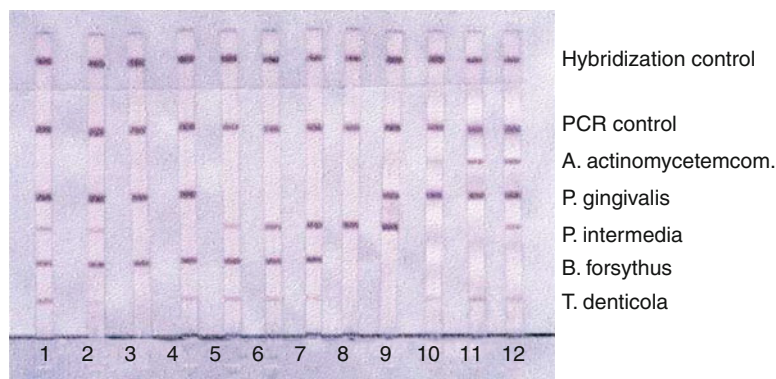
Eick and Pfister (2002) compared a commercial multiplex PCR of 16S rDNA for *A. actinomycetemcomitans*, *P. gingivalis*, *P. intermedia*, *T. forsythia*, and *T. denticola* with standard culturing. The PCR test was able to detect *P. gingivalis* and *T. forsythia* more often than cultivation. *A. actinomycetemcomitans* was detected in similar numbers with both techniques. Recently, the multiplex PCR molecular method for highly preserved regions 16S rDNAs was compared with the conventional culture method. The new technique showed a low sensitivity with higher values of specificity, with scarce possibilities therefore of false positives in the search for *A. actinomycetemcomitans*, *E. corrodens*, and *P. intermedia* (D’Ercole et al. 2008).

For some years, the DMDxA test (Omnigene, Cambridge, MA, USA) has been made commercially available. Only a few laboratories such as ANAWA (Wangen, Switzerland) are allowed to offer this test to the periodontist. Recently, Hain Diagnostika Ltd. (Nehren, Germany) has developed the microDentA kit, which can be used in each microbiological laboratory involved in the diagnosis of periodontopathogenic species. In this test, the multiplex PCR of 16S rDNA is followed by a simultaneous reverse hybridization for the species *A. actinomycetemcomitans*, *P. gingivalis*, *P. intermedia*, *B. forsythus*, and *Treponema denticola* (Fig. 12.39) (Eick and Pfister 2002).

12.4.8.3 Real-Time PCR

Real-time PCR, also referred to as quantitative PCR, quantitative reverse transcription-PCR, reverse transcription-quantitative PCR, and kinetic PCR, is a method used to quantify the copy numbers of DNA in clinical samples. There are two types of real-time PCR, namely, an intercalator-based method and a probe-based method. The intercalator-based method, also known as the SYBR Green method, intercalates SYBR Green, which binds to newly synthesized double-stranded DNA, producing a fluorescently labelled

Fig. 12.39 Reverse hybridization of subgingival plaque samples by the microDentA test (Eick and Pfister 2002. Reprinted with permission from John Wiley & Sons, Inc.)



PCR amplicon. The probe-based method, or *TaqMan* PCR, is more specific in that it utilizes a fluorogenic-labelled probe that binds only to its complementary sequence in the internal portion of the generated PCR amplicon (Paster and Dewhirst 2009).

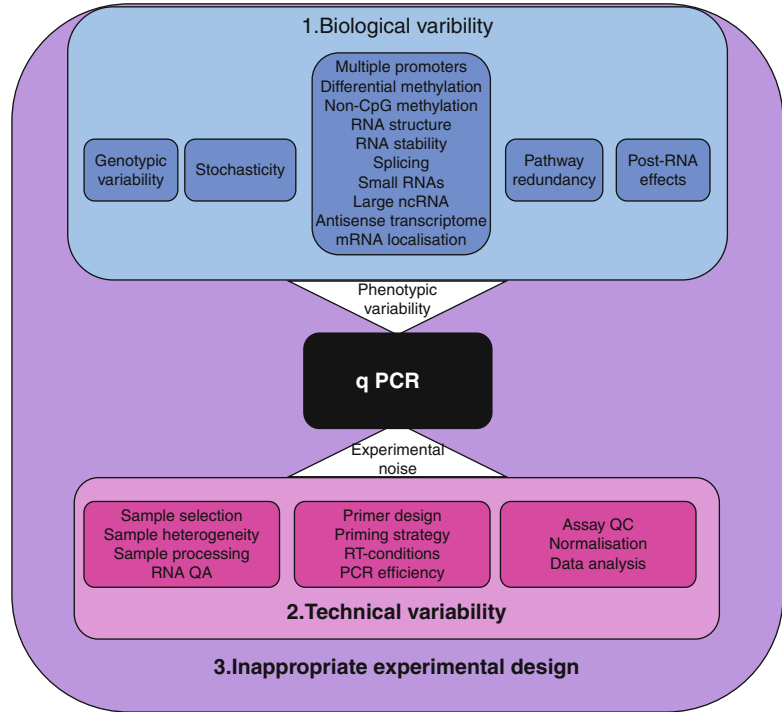
Real-time PCR has been used to detect and quantify several periodontal pathogens, including *A. actinomycetemcomitans*, *P. gingivalis*, *T. denticola*, *Fusobacterium nucleatum*, *T. forsythensis*, and *Prevotella intermedia*; the *tetQ* gene; and total bacteria, in clinical samples (Abiko et al. 2010; Asai et al. 2002; Boutaga et al. 2003, 2005, 2006, 2007; Braga et al. 2010; Casarin et al. 2010; Combs et al. 2008; Jervøe-Storm et al. 2005; Lau et al. 2004; Morillo et al. 2004; Sakamoto et al. 2001; Sakamoto et al. 2004; Yoshida et al. 2004).

MyPerioPath™ from OralDNA Labs™ (Brentwood, TN) is a commercially available service that utilizes *TaqMan* PCR to determine the presence and the microbial profile of 13 putative periodontal pathogens from oral specimens provided by clinicians. Treatment considerations and follow-up recommendations are given with the final report (Paster and Dewhirst 2009).

Several studies demonstrated a good harmony between quantitative PCR technology and the culture procedure (Boutaga et al. 2003, 2005, 2007; Lau et al. 2004; Jervøe-Storm et al. 2005; Nonnenmacher et al. 2005). A recent meta-analysis (Atieh 2008) evaluated the diagnostic accuracy of real-time PCR in detecting these two species, using the bacterial culture as a refer-

ence method. Overall, real-time PCR detected many more positive samples than cultivation because of its excellent detection limits compared to culture. This can be explained by culture's higher threshold of detection level (10^3 – 10^4 bacterial cells), thus limiting its sensitivity for a low number of bacteria. Moreover, bacterial culture is based on living cells, whereas in real-time PCR, the probes and/or primers detect the DNA of living and dead bacteria. Meta-analyses examined the areas under summary receiver operator characteristic curves (SROCs), and diagnostic odds ratios (DORs) were calculated to evaluate the test accuracy. The area under the SROC curve (AUC) is a global measure of overall test accuracy. An AUC of 100% indicates a perfect discriminatory ability. In Atieh (2008) meta-analysis, the AUC was close to 100%, with 95.9% for detecting *A. actinomycetemcomitans* and 99.5% for detecting *P. gingivalis*. The summary DORs in detecting *A. actinomycetemcomitans* and *P. gingivalis* were 18.5 and 40.47, respectively. These values indicate an overall high diagnostic accuracy of real-time PCR. However, the areas under the SROC curves and DORs obscure important details about test performance, particularly the differentiation between the detection of cases (sensitivity) and the identification of noncases (specificity). Real-time PCR has sensitivity ranging from 87% to 97% and specificity ranging from 67% to 77% for detecting *A. actinomycetemcomitans* and sensitivity ranging from 94% to 99% and specificity ranging from 55%

Fig. 12.40 Parameters affecting qPCR data and conclusions. Biological and technical variability are considerably higher for *in vivo* biopsies or environmentally sampled organisms compared to cells or organisms studied in the laboratory under standardised experimental conditions. Together with inappropriate experimental design they are the major confounding factors affecting pertinent conclusions of qPCR data (Bustin 2010b. Reprinted with permission from Elsevier)



to 68% for detecting *P. gingivalis*. Accordingly, real-time PCR is unlikely to miss any case, and a negative test is reliable for the exclusion of either species (Atieh 2008).

Real-time quantitative PCR (qPCR) has become the benchmark technology for the detection and quantification of nucleic acids in a research, diagnostic, forensic, and biotechnology setting (Bustin 2010a). Although PCR is an extremely sensitive technique, being able to detect even one copy of the DNA fragment, it has a number of important limitations. There are three obvious causes for the large number of variable, even contradictory, results obtained by PCR-based assays and published in the peer-reviewed scientific literature (Fig. 12.40). They are rather obvious and, one would like to think, well understood.

- **Biological variability:** The first cause is the evident reality that biology describes variability; hence, experiments will never yield identical results.

- **Technical variability:** The second reason derives from measurement errors that define technical variability. It describes the noise introduced into the assay and is inherent in any molecular technology.
- **Inappropriate experimental design:** The third explanation relates to inappropriate underlying assumptions generating results that exist in isolation, often are biologically or clinically of little consequence, may have negligible translational relevance, and frequently are wholly wrong (Bustin 2010).

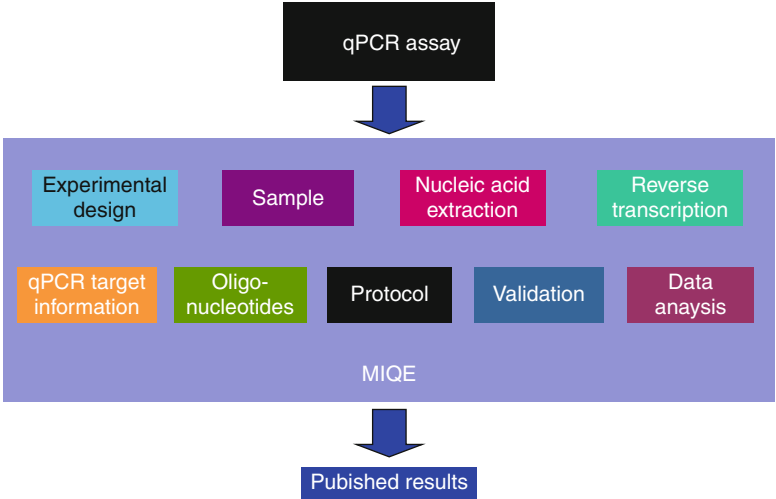
Proper experimental design is the key to any gene expression study. Since mRNA transcription can be sensitive to external stimuli that are unrelated to the processes studied, it is important to work under tightly controlled and well-defined conditions. Taking the time to define experimental procedures, control groups, type and number of replicates, experimental conditions, and sample handling methods within each group is essential to minimize variability (Table 12.3). Each of

Table 12.3 RT-qPCR experimental design and sample management (Taylor et al. 2010) (Reprinted with permission from Elsevier)

Experiment design	Control groups	Replicates	Experiment conditions	Sample handling
Disease or treatment groups	Time course study (i.e., $t=0$)	Biological (different sample per well)	Growth conditions (media and time or OD)	Precise time to harvest cells or tissue
Target genes implicated	Normal vs. disease (i.e., normal)	Technical (same sample per well)	Days of embryonic development	Sample extraction method
Potential reference genes	Untreated vs. drug treated (i.e., untreated)	–	Amount per mass of drug or compound	Preservation method and time
Number of data points to draw statistically significant conclusions	–	–	Sex, phenotype	Thaw and homogenization procedure
–	–	–	Incubation time	Total RNA extraction procedure

This table summarizes the workflow of a typical RT-qPCR experiment from experimental design to defining control groups, replicates, and experimental conditions to the detailed procedures for sample handling. This ultimately assures that the key steps in RT-qPCR data production lead to high quality, reproducible, and publishable data

Fig. 12.41 MIQE components. The qPCR assay is separated into nine major components that contain detailed information on pre- and post-assay parameters as well as comprehensive documentation of the experimental protocol (Bustin 2010b. Reprinted with permission from Elsevier)



these parameters should be carefully recorded prior to conducting gene expression experiments to assure good biological reproducibility for published data (Taylor et al. 2010).

These issues are being addressed by a set of guidelines that propose a minimum standard for the provision of information for qPCR experiments (‘MIQE’) (Fig. 12.41). MIQE aims to restructure today’s free-for-all qPCR methods into a more consistent format that will encourage detailed auditing of experimental detail, data analysis, and reporting principles. General imple-

mentation of these guidelines is an important requisite for the maturing of qPCR into a robust, accurate, and reliable nucleic acid quantification technology (Bustin 2010b).

12.4.9 Terminal Restriction Fragment Length Polymorphism (T-RFLP)

Pattern-based technologies, such as restriction fragment length polymorphism analysis (Allaker et al. 1997; Beikler et al. 2003a, b; Loos and Dyer

1992; Loos et al. 1990; Riviere et al. 1999; Sato and Kuramitsu 1999), multi-locus enzyme electrophoresis (Loos et al. 1993; Rumpf et al. 2000), arbitrarily primed PCR for randomly amplified polymorphic DNA analysis (Alpha et al. 2001; Asikainen and Chen 1999; Asikainen et al. 1995; Dogan et al. 1999a, b; Ehmke et al. 1999; George et al. 1997; Huang et al. 2003; Lakio et al. 2002; Menard et al. 1992, 1994; Menard and Mouton 1993, 1995; Riviere et al. 1999; Correia et al. 2004; Perez-Chaparro et al. 2008, 2009), and amplified fragment length polymorphism analysis (Rijnsburger et al. 2007; Yoshino et al. 2007a, b), have been used for characterizing periodontal pathogens (Kuboniwa et al. 2010). However, it should be noted that patterns produced in different laboratories can be compared only if very strict quality standards are followed (Kuboniwa et al. 2010).

Terminal restriction fragment length polymorphism (T-RFLP) analysis is a popular high-throughput fingerprinting technique used to monitor changes in the structure and composition of microbial communities. T-RFLP analysis is one of the most frequently used high-throughput fingerprinting methods. Because of its relative simplicity, T-RFLP analysis has been applied to the analysis of fungal ribosomal genes, bacterial 16S rRNA genes, and archaeal 16S rRNA genes. In addition, T-RFLP has been used for the analysis of functional genes such as those encoding for nitrogen fixation and methane oxidation. However, most frequently, the technique is used to amplify small subunit (16S or 18S) rRNA genes from total community DNA using polymerase chain reaction (PCR) wherein one or both of the primers used are labelled with a fluorescent dye. The resulting mixture of rRNA gene amplicons is then digested with one or more restriction enzymes that have four base-pair recognition sites, and the sizes and relative abundances of the fluorescently labelled T-RFs are determined using an automated DNA sequencer. Since differences in the sizes of T-RFs reflect differences in the sequences of 16S rRNA genes (i.e. sequence polymorphisms), phylogenetically distinct populations of organisms can be resolved. Thus, the pattern of T-RFs is a composite of DNA frag-

ments with unique lengths that reflects the composition of the numerically dominant populations in the community. While T-RFLP shares problems inherent to any PCR-based method, it has been shown to provide a facile means to assess changes in microbial community structure on temporal and spatial scales by monitoring the gain or loss of specific fragments from the profiles. When coupled with 16S rRNA clone library construction and clone sequencing, additional specific information on the composition of microbial communities can be obtained (Schütte et al. 2008). In dental field, the T-RFLP analysis has been applied to assess oral microbial communities in human saliva (Fig. 12.42) (Sakamoto et al. 2003; Takeshita et al. 2009). Sakamoto et al. (2004) used T-RFLP analysis to study the change of oral microbiota in saliva and subgingival plaque samples of patients with periodontitis before and 3 months after periodontal treatment. Significant changes in the T-RFLP patterns of subgingival plaque samples of the patients were noted after 3 months of improved oral hygiene and full-mouth supra- and subgingival scaling and root planing. Although the proportions of T-RFs larger than 1,000 bp were notable in the T-RFLP patterns generated after digestion with HhaI of the samples from the patients before treatment, the proportions of these T-RFs were significantly reduced or not detected after treatment (Sakamoto et al. 2005).

12.4.10 Problems Associated with Microbial Testing

A problem that is frequently brought up in microbial testing is the potential for differences in the results obtained by different laboratories using different techniques (Shaddox and Walker 2009).

A field study using five different private periodontal practices was conducted by Mellado et al. (2001); it compared two microbiologic culture samples simultaneously secured from the same sites within 23 individual patients and submitted for bacterial identification and antibiotic sensitivity testing to two separate laboratories.

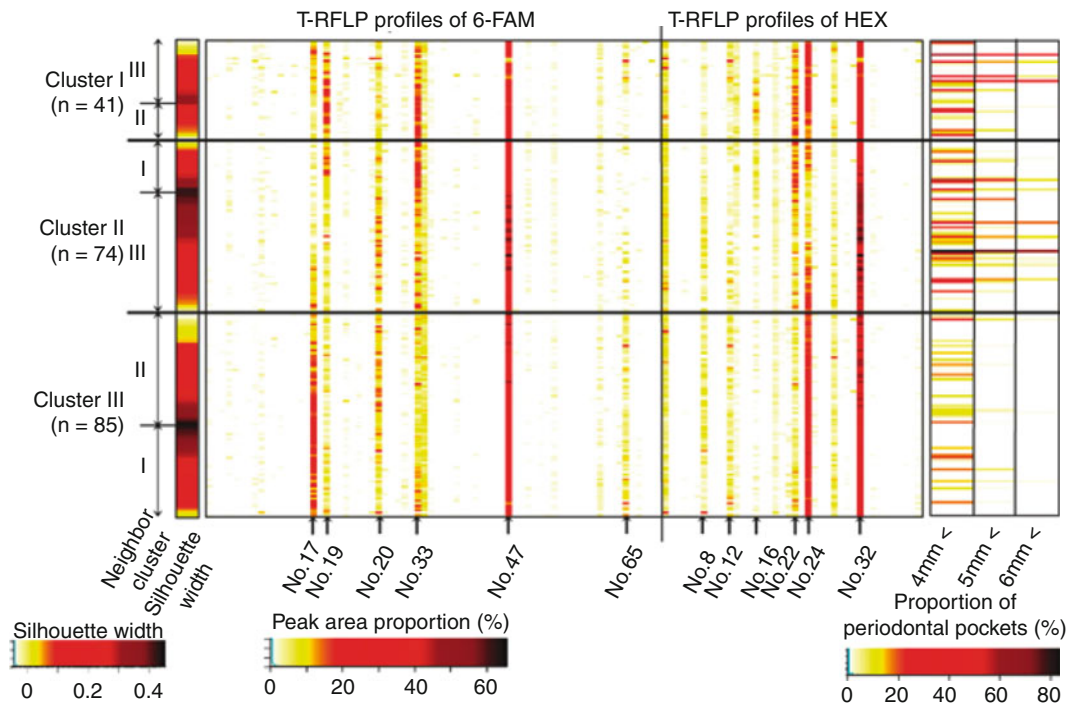


Fig. 12.42 Terminal restriction fragment length polymorphism (*T-RFLP*) peak patterns from 200 subjects (rows) visualized as a gel-like image. The patterns were sorted into three clusters, separated based on their nearest neighbor cluster in each cluster, and ordered according to each silhouette width, which measures the degree of fitness of an object to its cluster, within the cluster. The silhouette width of each sample (displayed as *gray scale bars*) and the nearest neighbor cluster are indicated to the left of the image. The area proportion in each T-RFLP

profile (6-carboxyfluorescein (6-FAM) and hexachlorofluorescein (HEX)) of individual peaks (69 detected 6-FAM peaks and 41 detected HEX peaks; columns) is represented as the *gray scale* intensity in each grid. The percentage of sites with periodontal pockets (pocket depths over 4, 5 or 6 mm, respectively) in all surveyed sites of each subject are displayed as *gray scale bars* to the right of the image (Takeshita et al. 2009. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

The results from the two laboratories were often different. In no instance did both laboratories agree on the presence of identical bacterial species. When only bacteria above threshold levels were compared, agreement was found in only nine of 23 cases. When examining antibiotic sensitivity, using 100% kill of all tested pathogens as the ideal, agreement between the two laboratories was poor. The laboratories agreed on the use of amoxicillin 17% of the time, tetracycline 26% of the time, and metronidazole 48% of the time. The use of amoxicillin and metronidazole in combination yielded a 78% agreement when the results of both laboratories were combined (Mellado et al. 2001). A similar study by the same group,

2 years later, using four different private periodontal practices and two independent laboratories was conducted to compare two microbiologic cultures sampled simultaneously from the same sites in 20 individual patients. Both paired samples were submitted separately to one of the two independent laboratories for bacterial identification and antibiotic sensitivity testing. The results from the two samples were quite variable. In only two instances did both specimens reveal the presence of identical bacterial species, but these specimens differed in both threshold levels and antibiotic sensitivity. When only bacteria above threshold levels were compared, total agreement was found in 11 of 20 cases. When

examining antibiotic sensitivity, using 100% kill as the ideal, agreement between the two specimens was inconsistent. The use or nonuse of tetracycline was in agreement 85% of the time, amoxicillin 75% of the time, metronidazole 70% of the time, and amoxicillin–metronidazole in combination 85% of the time. The two specimens agreed on the empirical use of amoxicillin 45% of the time, tetracycline 60% of the time, and metronidazole 60% of the time. The empirical use of amoxicillin–metronidazole in combination yielded 80% agreement when the results of both specimens were combined. This supports the data from a previous study that examined specimens secured simultaneously from the same site and submitted to two different testing laboratories. The failure of microbial testing to achieve a higher level of consistency between samples leaves the clinical efficacy of microbial testing in question (Salkin et al. 2003). In contrast, a study performed by Cohen et al. (2001) found that for these three testing laboratories, the results were very similar microbiologically (*P. gingivalis*, *F. nucleatum* and *C. rectus*) and the differences that were present would not influence the antibiotic selected.

Other problems frequently associated with microbial testing are related to the clinical staff collecting the samples. These problems include contamination of the sample, failure to collect an adequate size of sample, pooling samples from a number of different pockets, and inadequacy of sample handling and transport. Any of these will affect the results obtained by the testing laboratory and can lead to the inappropriate selection of an antibiotic. Therefore, detailed instructions must be supplied to the clinical staff on proper sampling technique and handling and the importance of following these instructions (Shaddox and Walker 2009).

12.5 Biochemical Analysis as Part of Periodontal Diagnosis

The biochemical assessment of periodontal disease can be accomplished using several approaches. The most practical and least-invasive

approach involves analysis of biologic fluids that are derived from the periodontal tissues or contain specific mediators that are present as a result of periodontal disease. The biologic fluids that have been studied to understand the nature of destructive periodontitis and to identify potential diagnostic markers of active disease include serum (blood), gingival fluid, and saliva (Wolf and Lamster 2011).

Biomarkers have the potential to provide additional information over standard clinical indices. A periodontal disease biomarker should objectively measure the disease processes involved in the disease pathogenesis and reflect the underlying aetiology in view of ultimately personalized periodontal therapy. In practice, there is no such thing as a perfect biomarker; there always is a trade-off between specificity and sensitivity. The reason for doing the test will dictate both the choice of biomarker and the need to maximize sensitivity or specificity (Kinane et al. 2011).

Biomarkers might be useful at the patient as well as at the site level. For site-specific assessment, biomarkers should be collected from GCF. For patient-specific assessment, biomarkers should be sampled from peripheral blood, saliva, and pooled GCF from multiple sites (Kinane et al. 2011).

Periodontal disease biomarkers can be grouped as listed below (Kinane et al. 2011):

1. *Susceptibility*. A biomarker that prospectively identifies individuals or sites at increased risk for periodontal disease.
2. *Diagnostic*. A biomarker that identifies the presence of periodontal disease.
3. *Prognostic*. A biomarker that identifies patients or sites most likely to respond to specific interventions.
4. *Predictive*. A biomarker that predicts future progression of disease.
5. *Therapeutic*. A biomarker that provides a quantifiable measure of response to periodontal therapy.

All five categories of biomarkers might be useful at the patient as well as at the site level.

Are there any promising periodontal disease-soluble biomarkers to detect? (Kinane et al. 2011).

- (a) Disease susceptibility: We do not currently have reliable biomarkers of susceptibility.
- (b) Diagnosis: Most biomarkers reflect inflammation so their ability to discriminate between disease states is limited. There is limited evidence that salivary concentrations of MMP-8 and myeloperoxidase can distinguish between gingivitis and periodontitis. In cross-sectional studies comparing periodontitis and gingivitis patients, differences in the concentration or active states of specific biomarkers can be demonstrated. Because periodontitis differs from gingivitis in terms of bone destruction, biomarkers specific for bone loss such as RANKL/osteoprotegerin ratio (RANKL/OPG) may be more suited to differentiating gingivitis and periodontitis. Analysis of GCF samples for a range of possible markers including prostaglandin E₂, β glucuronidase, oncostatin M, cathepsin B and K, and carboxyterminal telopeptide pyridinoline cross-links of type I collagen (ICPT) may have some diagnostic utility for periodontal disease. There is currently no reliable biomarker to differentiate CP from aggressive periodontitis.
- (c) Prognosis: No existing biomarkers have been proven to have prognostic utility either at the site or patient level.
- (d) Predictive: There is limited evidence that certain molecules may be useful biomarkers of future periodontal disease progression.
- (e) Therapeutic: Many inflammatory markers have been shown to decrease in response to periodontal treatment.

12.5.1 Source for Samples of the Different Diagnostic Tests: Blood Components

Periodontitis is bacterial infection of tooth-supporting tissues leading to inflammation and, subsequently, to loss of teeth. It is one of the most common infections worldwide. Recent studies have shown that periodontal infection may pose a threat to general health by increasing the risk of cardiovascular and lung diseases

and preterm labour. Thus, useful markers of systemic exposure to periodontitis are needed. The measurable markers of periodontitis in serum can be divided in two types: ones that derive directly from periodontopathic pathogens and the others that originate in the defence mechanisms by the host. Viable periodontal pathogens and their DNA have been detected from human atherosclerotic lesions but not in blood circulation, probably due to the methodological difficulties. Periodontitis is associated with endotoxaemia, which has been shown indirectly as elevated concentrations of lipopolysaccharide (LPS) binding protein, soluble CD14, and antibodies to LPS of periodontal pathogens, but also directly as elevated concentrations of endotoxin in patients with periodontitis compared with periodontally healthy subjects. No method for determining serum concentrations of specific LPS of periodontopathic pathogens, however, has been reported (Pussinen et al. 2007).

Most of the studies using serum markers of periodontitis have concentrated on measuring the host response against the bacterial insult (**Fig. 12.1**). These include matrix metalloproteinases, a variety of cytokines and chemokines, inflammation markers, antiphospholipid antibodies, and—most commonly—antibodies to periodontal pathogens. From these, only antibodies to periodontopathic pathogens can be considered a specific marker of systemic exposure, even though they do not necessarily display disease status, activity, or prognosis. Since there are several periodontopathic pathogens and as they exhibit intraspecies differences, the choice of the antigen in the assay arsenal is crucial. Although no laboratory determination can replace clinical periodontal examination, availability of such a method would be useful in large-scale studies (Pussinen et al. 2007).

Various studies have evaluated the molecular markers of tissue destruction in serum or plasma: These manifestations of periodontal diseases are mainly sought to clarify the possible interactions between periodontitis and various systemic diseases and/conditions such as cardiovascular diseases (CVDs), pregnancy complications, diabetes mellitus, and rheumatoid arthritis. Serum or plasma provides information

about the inflammatory stimulus and/or response generated in circulation towards the periodontal pathogens that colonize in the subgingival area. Data from the selected studies evaluating the potential utility of serum components as diagnostic markers for periodontal tissue destruction are presented in **Table 12.4** (Buduneli and Kinane 2011).

12.5.2 Source for Samples of the Different Diagnostic Tests: Saliva Components

Saliva is the fluid that bathes the hard and soft tissues of the oral cavity. Human saliva is produced by three major glands—the parotid, submandibular, and sublingual—as well as by numerous minor glands (Sahingur and Cohen 2004).

12.5.2.1 Types of Human Saliva

Saliva can be assessed as whole (mixed) or as gland-specific saliva. The composition of saliva is affected by many factors, such as the originating gland, diet, use of pharmacological agents, and the systemic or oral health of the patient. Parotid saliva is a serous, watery secretion whereas the submandibular glands produce a mixed serous and mucous fluid; a predominantly mucous fluid is derived from the sublingual glands.

Whole saliva consists of a mixture of oral fluids and includes secretions of the major and minor salivary glands, in addition to constituents of non-salivary origin derived from GCF, expectorated bronchial secretions, serum, and blood cells from oral wounds, as well as bacteria and bacterial products, viruses and fungi, desquamated epithelial cells, and food debris (**Fig. 12.43**). Factors derived from GCF and the subgingival plaque are found in whole saliva rather than gland-specific saliva. In reviewing the available data, analysis of whole saliva holds greater promise than gland-specific saliva when considering the development of a diagnostic test for periodontal disease. This conclusion is based on the apparent importance of factors derived from GCF which are present in whole saliva (Kaufman and Lamster 2000). Some of the components

commonly found in mixed human saliva are noted in **Table 12.5** (Sahingur and Cohen 2004).

12.5.2.2 Techniques Used to Collect and Analyze Saliva

Both whole and gland-specific saliva can be collected with or without gustatory stimulation. Unstimulated whole saliva is composed of secretions from the parotid, submandibular, sublingual, and minor mucous glands, as well as from gingival crevicular fluid, desquamated epithelial cells, microorganisms, leukocytes, food residue, and blood. Stimulated saliva has been obtained in response to masticatory or gustatory stimulation using a variety of methods, including paraffin wax, rubber bands, gum base, and citric acid. Some of the more common techniques for obtaining whole- and duct-collected saliva are described in **Table 12.6** (Sahingur and Cohen 2004).

Since saliva can be collected more easily and less invasively compared to blood, interest in saliva as a diagnostic fluid has received increased attention during the past two decades. Indeed, saliva has been used as a diagnostic fluid in medicine, where clinical problems such as digitalis toxicity (calcium and potassium measurement), affective disorders (prostaglandins), immunodeficiency (secretory IgA), and tobacco usage (cotinine) have been assessed (Sahingur and Cohen 2004).

12.5.2.3 Salivary Markers for Periodontal Diagnosis

The use of saliva for periodontal diagnosis has been the subject of considerable research activity, and proposed markers for disease include proteins of host origin (i.e. enzymes, immunoglobulins), phenotypic markers (epithelial keratins), host cells, hormones (cortisol), bacteria and bacterial products, volatile compounds, and ions (Kaufman and Lamster 2000; Sahingur and Cohen 2004; Ozmeric 2004; Taba et al. 2005; Kinney et al. 2007; Zhang et al. 2009; Giannobile et al. 2009; Gorr 2009; Gorr and Abdolhosseini 2011; Saygun et al. 2011) (**Table 12.7**).

Virulence factors from periodontopathic bacteria either cause degradation of host tissue directly or activate a host response. The latter

Table 12.4 Studies evaluating possible biomarkers in serum/plasma samples (Buduneli and Kinane 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Al-Ghamdi and Anil (2007)	Cross-sectional	30 smoker CP, 30 non-smoker CP, 30 healthy			Serum	Immunoglobulins	Smoking decreases Ig content
Amarasena et al. (2008)	Follow-up	266 elderly	6 years Progression: CAL increase >3 mm	PD, CAL	Serum	Albumin, Ca, IgG, A, M	Serum Ca correlated with PD, may be a risk factor for progression
Behle et al. (2009)	Intervention	30 CP	Before, 4-weeks after SRP		Plasma	19 biomarkers, multiplex	Overall reduction in systemic inflammation, great personal variation
De Queiroz et al. (2008)	Cross-sectional	17 CP, 8 healthy			Serum	24 cytokines with multiplex	RANTES differed between groups
Duerte et al. (2010)	Intervention	14 healthy controls, 14 CP, 14 GAgP	At baseline and 6 months after SRP		Serum	IL-17, IL-4, IL-23, IFN- γ , TNF- α	IL-17, TNF- α levels were higher in GAgP than CP, healthy both at baseline and after SRP
Furugen et al. (2008a, b)	Cross-sectional	158 Japanese elders			Serum	Adiponectin, resistin, IL-6, TNF- α	Increased serum cortisol is associated with BOP
Glas et al. (2008)	Cross-sectional	105 CP, 122 healthy			Plasma	Surfactant protein D	Increased in CP; may be a biomarker for CP
Guentisch et al. (2009)	Cross-sectional	425 non-smoker, systemically healthy adults		Number of teeth with PD > 4–6 mm	Serum	IL-6, TNF- α	Serum IL-6 is associated with periodontal disease severity
Ishisaka et al. (2008)	Cross-sectional	467 adults		PD, CAL, BOP	Serum	Cortisol	Serum cortisol associated with PD, CAL

(continued)

Table 12.4 (continued)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Iwasaki et al. (2008)	Follow-up	600 elderly	4 years	CAL	Serum	Albumin	Serum albumin may be a risk predictor for progression
Liu et al. (2010)	Cross-sectional	40 CP, 40 healthy		PD, CAL, BOP	Serum	CRP, lipid profile	CRP higher in CP, HDL lower in CP. Correlated with clinical parameters
Marcaccini et al. (2009)	Intervention	25 CP, 20 healthy	3 months after SRP	PD, BOP, CAL	Serum, plasma	hsCRP, IL-6, CD40 ligand, MCP-1, sP-selectin, sVCAM-1, sICAM-1	SRP decreased CRP and IL-6, others were not affected
Nakajima et al. (2010)	Intervention	78 CP, 40 healthy	At baseline, after SRP		Serum	hsCRP, IL-6, TNF- α	CRP, IL-6 higher and TNF lower in CP, IL-6, CRP decreased after SRP
Nibali et al. (2007)	Cross-sectional	302 CP, 183 healthy			Serum	Inflammatory, metabolic markers	CP may be related with systemic inflammation, dysmetabolic state
Nicu et al. (2009)	Cross-sectional	105 CP, 57 healthy	6 months	PD, CAL, BOP, subgingival calculus	Serum	CRP, sCD14	sCD14 increased in CP
Offenbacher et al. (2009)	RCT	151 SRP, 152 community care			Serum	CRP	Serum CRP was not affected by SRP
Pradeep et al. (2010)	Intervention	20 healthy, 20 gingivitis, 20 CP	CP patients also 6–8 weeks after SRP	PD, CAL, GI	Serum	Oncostatin M	Highest in CP, decreased after SRP, not detectable in gingivitis, healthy. Correlated with PD, CAL
	Cross-sectional	52 subjects		PD, CAL, BL	Mouth-wash, serum	IL-6	Serum IL-6 increased with periodontal disease severity

Raunio et al. (2009)	Cross-sectional	56 CP, 28 healthy			Serum	sCD14	Higher in CP, correlated with periodontal disease extent
Renvert et al. (2009)	RCT	28 CP test (anti-inflammatory therapy), 29 CP placebo	PD, BOP		Serum	hsCRP, IL-6, IL-1 β , IL-8, IL-12, TNF- α , IFN- γ	CRP, IFN, IL-6 decreased, others did not change
Trindade et al. (2008)	Cross-sectional	29 CP, 12 AgP, 22 gingivitis, 26 healthy			Serum	Antibody levels against <i>P. gingivalis</i>	Serum Ig titres correlated with clinical periodontal diagnosis
Yoshihara et al. (2009)	Cross-sectional	148 subjects at age 77	CAL		Serum	Osteocalcin, bone-specific alkaline phosphatase	Negative correlation between CAL > 6 mm and osteocalcin

BL bone level, *CP* chronic periodontitis, *CAL* clinical attachment level, *PD* probing depth, *Ig* immunoglobulin, *RANTES* regulated upon activation normal T-cell expressed and secreted, *GAgP* generalized aggressive periodontitis, *SRP* scaling and root planing, *IL* interleukin, *IFN- γ* interferon- γ , *TNF* tumour necrosis factor, *BOP* bleeding on probing, *CRP*, C-reactive protein, *HDL* high-density lipoprotein, *MCP* monocyte chemotactic protein, *sVCAM* soluble vascular cell adhesion molecule, *sICAM* soluble intercellular adhesion molecule

Fig. 12.43 Components of whole saliva (Kaufman and Lamster 2000. Reprinted with permission from John Wiley & Sons, Inc.)

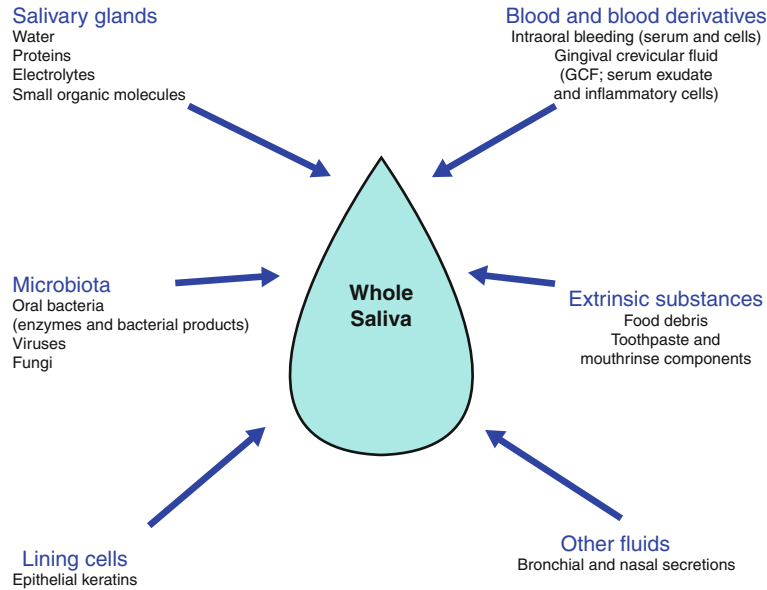


Table 12.5 Selected components of human mixed saliva (Sahingur and Cohen 2004) (Reprinted with permission from John Wiley & Sons, Inc.)

Electrolytes	Organic compounds	Proteins and macromolecules
Ammonia	Amino acids	Aggregins
Bicarbonate	Creatinine	Albumin
Calcium	Fatty acids	Blood group substances
Chloride	Glucose	Cystatins
Fluoride	Lactic acid	Enzymes
Hypothiocyanate	Lipids	Fibronectin
Iodine	Sialic acid	Fucose-rich glycoprotein
Magnesium	Urea	Growth factors
Nitrate	Uric acid	Gustin
Nitrite		Histatins
Phosphates		Immunoglobulins
Potassium		Kallikrein
Sodium		Lactoferrin
Sulfate		Mucin glycoproteins
Thiocyanate		Proline-rich proteins
		Ribonucleases
		Secretory component
		Serum proteins
		Tyrosine-rich proteins

initiates the release of biological mediators from host cells and, when exaggerated in nature, leads to host tissue destruction. These mediators include proteinases, cytokines, and prostaglandins. Various bacteria-derived enzymes, such as collagen-degrading enzymes, elastase-like enzymes, trypsin-like proteases, aminopeptidases, and dipeptidylpeptidases, are recognized as important participants in tissue destruction. Host- and bacteria-derived enzymes, proteins, and other inflammatory mediators appear to hold great promise as salivary biomarkers for the diagnosis of periodontal disease. Specific salivary proteomic biomarkers have been identified for three key features of the pathogenic processes in periodontal disease—**inflammation, collagen degradation, and bone turnover (Fig. 12.44)**. Innate host defence responses are triggered by bacterial lipopolysaccharide and other microbial components and products (e.g. bacterial DNA). As a result, neutrophilic polymorphonuclear leukocytes, monocytes, and activated macrophages are recruited to the site and release numerous cytokines, such as prostaglandin E₂, tumour necrosis factor (TNF), and interleukins IL-1 and IL-6, which direct further inflammatory processes.

Table 12.6 Saliva collection methods (Sahingur and Cohen 2004) (Reprinted with permission from John Wiley & Sons, Inc.)

Saliva	Method
Mixed	<i>Draining/spitting method:</i> The subject is asked to accumulate saliva in the floor of the mouth and then spit into a pre-weighed or graduated test tube (Atkinson et al. 1993; Navazesh 1993)
Mixed	<i>Suction method:</i> Saliva is continuously aspirated from the floor of the mouth into a suitable collection vessel (Atkinson et al. 1993; Navazesh 1993).
Mixed	<i>Swab (absorbent) method:</i> A pre-weighed swab, cotton roll, or gauze sponge is placed in the mouth at the orifices of the major glands and is removed for reweighing at the end of the collection period (Atkinson et al. 1993; Navazesh 1993).
Parotid	Modified Carlson-Crittenden device. The device has two chambers, an inner and an outer chamber. The inner chamber is placed over the orifice of the parotid duct and attached to tubing that carries saliva to the collection vessel. The outer chamber is connected to a vacuum squeeze bulb. The bulb is compressed and the collector is then placed over the duct opening (Shannon and Chauncey 1967)
Submandibular/sublingual	Since submandibular and sublingual secretions enter the oral cavity via a common duct system, it is difficult to collect individual salivas separately. For combined submandibular-sublingual saliva, custom-made collectors can be fabricated. Those devices contain a central chamber for the collection of submandibular saliva and one or two side chambers for the collection of sublingual saliva (Parr and Bustos-Valdes 1984).
Submandibular/sublingual	Alternatively, Fox (1989) have described a simpler method for the collection of submandibular-sublingual. After blocking the parotid saliva secretion by placing a gauze pad at the orifice of the parotid ducts, saliva can be collected from the floor of the mouth with a micropipette.
Minor glands	Minor gland secretions can be collected by micropipette, absorbent filter paper or strips from the inner surface of lips, palate, or buccal mucosa and quantitated by weight differences or using a Periotron® device (Shern et al. 1990).

As a consequence, matrix metalloproteinases (MMPs), which are powerful collagen-destroying enzymes, are produced by alveolar bone and polymorphonuclear leukocytes. Pyridinoline cross-linked carboxyterminal telopeptide and osteocalcin are then released into the surrounding area and transported through gingival crevicular fluid into the periodontal pocket (Zhang et al. 2009).

Salivary enzymes originate from three major sources: the actual salivary secretions, the host cells found in GCF, and finally disposed bacterial cells from dental plaque and mucosal surfaces. Data from the selected studies evaluating the potential utility of salivary components as diagnostic markers for periodontal tissue destruction are presented in **Table 12.8** (Buduneli and Kinane 2011).

One obstacle associated with many studies reporting associations between potential salivary markers and periodontal diseases is relatively small study populations. Considerable variations

in the concentrations of salivary components also are routinely observed in individuals and populations, among healthy patients as well as those with periodontal diseases. Consequently, few well-accepted normal values have been proposed for any specific marker, deviations from which might be considered abnormal. Moreover, most authors have used cross-sectional designs to analyse salivary components of interest that have arisen following the development of periodontal diseases. Although such factors may represent markers having potential diagnostic value, they may not necessarily provide information regarding a particular individual's susceptibility to periodontitis prior to the onset of periodontal breakdown (Sahingur and Cohen 2004).

12.5.2.4 Saliva Ontology

The rapid development and maturity of the genomics field have led to the emergence of other 'omics' studies, such as proteomics, tran-

Table 12.7 Salivary biomarkers for periodontal disease (Zhang et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Proteomic biomarkers		Genetic biomarkers	Microbial biomarkers	Other biomarkers
Alpha-glucosidase	IgM	Cathepsin C gene mutation	<i>Aggregatibacter actinomycetemcomitans</i>	Calcium
Acid phosphatase	Kallikrein	Collagen gene mutation		Cortisol
Alkaline phosphatase	Kininase	IL-1 polymorphisms	<i>Campylobacter rectus</i>	Hydrogen sulfide
Aminopeptidases	Lactoferrin	IL-10 polymorphisms	Mycoplasmas	Methylmercaptan
Aspartate aminotransferase	Lactotransferrin	Tumor necrosis factor polymorphisms	<i>Porphyromonas gingivalis</i>	Picolines
Beta-galactosidase	Lactate dehydrogenase		<i>Prevotella intermedia</i>	Polymorphonuclear leukocytes
Beta-glucosidase	Lysozyme		<i>Peptostreptococcus micros</i>	Pyridine
Beta-glucuronidase	MMP-13		<i>Prevotella nigrescens</i>	
Calprotectin	MMP-8		<i>Treponema denticola</i>	
Caprylate esterase	MMP-9		<i>Tannerella forsythia</i>	
Cathepsin B	Myeloperoxidase		<i>Treponema socranskii</i>	
CD14	Osteocalcin			
Cystatins	Osteonectin			
Elastase	Osteopontin			
Epidermal growth factor	Platelet-activating factor			
Esterase	Platelet-derived growth factor			
Fibronectin	Pyridinoline cross-linked carboxyterminal telopeptide			
Gelatinase	sIgA (secretory IgA)			
IgA	Trypsin			
IgG	Vascular endothelial growth factor			

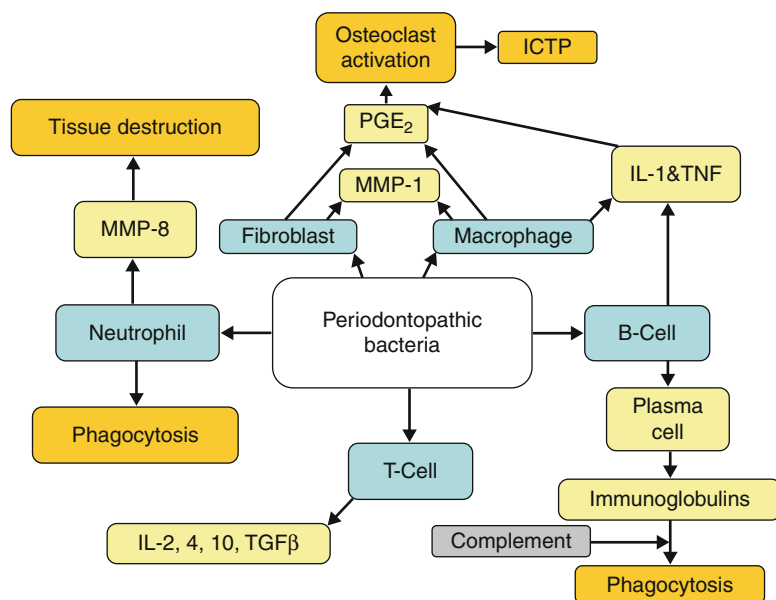
Fig. 12.44 Flow chart indicating cells, enzymes, immunoglobulins and other potential proteomic biomarkers for periodontal disease and their involvement in its pathogenesis. *MMP* matrix metalloproteinase, *IL* interleukin, *TGF* transforming growth factor, *TNF* tumor necrosis factor, *PGE₂* prostaglandin *E₂*, *ICTP* pyridinoline cross-linked carboxyterminal telopeptide (Zhang et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

Table 12.8 Studies evaluating possible biomarkers in saliva samples (Buduneli and Kinane 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Aemaimanan et al. (2009)	Cross-sectional	30 localized CP, 30 generalized CP, 30 controls		PD, CAL, BOP	Saliva, subgingival plaque	Alanine aminopeptidase Dipeptidyl peptidase IV, <i>P. gingivalis</i>	Dipeptidyl peptidase IV but not alanine aminopeptidase activity is associated with CP, presence of <i>P. gingivalis</i>
Aurer et al. (1999)	Cross-sectional	20 RPP, AP,			Saliva	IL-6	IL-6 increased in RPP, CRP decreased in AP
Baron et al. (1999)	Cross-sectional	8 PD patients, 16 controls			Saliva	Cystatins, total cystatin activity	Total cystatin activity decreased in patients
Cutundo et al. (2006)	Cross-sectional	37 patients, variable PD		CPI	Whole saliva	Melatonin	Inverse relation between salivary melatonin and PD severity
De la Pena et al. (2007)	Cross-sectional	175 volunteers		Number of teeth, PD	Saliva	Lactate dehydrogenase	Increased in periodontitis, correlated with PD, may be a biomarker
Diab-Ladki et al. (2003)	Cross-sectional	17 PD, 20 controls			Saliva	TAC, total protein content	TAC decreased, an imbalance between oxidants and antioxidants
Garito et al. (1995)	Cross-sectional	69 Periodontal disease patients		PD	Saliva	PAF	Salivary PAF correlated with PD severity
Ghallab and Shaker (2010)	Intervention	22 CP, 22 healthy (half smoker, half non-smoker)	At baseline, 1 month after SRP	PI, GI, PD, CAL only in CP	Saliva	sCD44	Highest in smoker CP, decreased after SRP in both smokers and non-smokers
Gheren et al. (2008)	Prospective	18 CP, 18 controls	30 days after perio. therapy	PD, CAL, PI, GI	Saliva	Salivary arginase activity	Salivary arginase activity increased in CP, candidate after therapy; a candidate salivary marker of periodontal status
Gürsoy et al. (2009)	Cross-sectional	84 CP, 81 controls			Saliva	Elastase, lactate dehydrogenase IL-6, IL-1 β , TNF- α	Only IL-1 β can differentiate periodontitis
Haigh et al. (2010)	Cross-sectional, intervention	Severe periodontitis			Saliva	Proteomics 128 proteins	S100 proteins highly related with PD severity

(continued)

Table 12.8 (continued)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Lamster et al. (2003)	Cross-sectional	380 subjects		PD, CAL, GI	Saliva, blood	β glucuronidase	Periodontal clinical parameters correlated with salivary β glucuronidase
Koss et al. (2009)	Cross-sectional	89 mild, moderate, severe CP 29 controls			Whole saliva	Peroxidase, hydroxyproline, sIgA	Peroxidase, hydroxyproline increased, sIgA diminished in CP
Nomura et al. (2006)	Cross-sectional	187 subjects			Saliva	Lactate dehydrogenase, AST, blood urea nitrogen	Salivary lactate dehydrogenase, AST, blood urea nitrogen have potential to screen PD
Özmeric et al. (2000)	Cross-sectional	20 CP, 15 controls		PD, CAL, PI, GI	Saliva	Total protein, arginase	Salivary arginase increased in CP
Su et al. (2009)	Cross-sectional	58 periodontitis 234 controls		CPTN	Whole saliva	ROS, TAC	ROS increased in periodontitis, TAC correlates with disease severity
Takane et al. (2002)	Cross-sectional	78 CP, 17 healthy			Saliva	8-OHdG	Higher in CP, can be a diagnostic marker
Teles et al. (2009)	Cross-sectional	74 CP, 44 healthy			Saliva multiplex	G-MCSF, IL-1 β , -2, -4, -5, -6, -8, -10 IFN- γ , TNF- α	No difference between groups. IL-8 correlated with PD, BOP, these cytokines cannot discriminate CP and healthy
Totan et al. (2006)	Cross-sectional	Periodontitis, controls		PD, BOP	Saliva	AST, ALP, aminopeptidases, glucuronidases	Salivary AST, ALP increased with pockets, BOP, suppuration
Wilczynska-Borawska et al. (2006)	Cross-sectional	26 PD, 20 healthy controls		GI, PBI, PI, PD, CAL, number of teeth	Whole saliva	Hepatocyte growth factor	Salivary HGF may be related with severity of PD
Yoshie et al. (2007)	Intervention	49 CP	4 weeks after SRP	PD, CAL, BOP	Saliva	AST, ALT, LDH	All decreased after SRP, they can be useful markers for inflammation, destruction

CP chronic periodontitis, PD probing depth, CAL clinical attachment level, BOP bleeding on probing, RPP rapidly progressive periodontitis, AP adult periodontitis, IL interleukin, CRP C-reactive protein, CPI community periodontal index, TAC total antioxidant capacity, PAF platelet activating factor, PI plaque index, GI gingival index, sCD44 soluble CD44, SRP scaling and root planing, TNF tumour necrosis factor, sIgA secretory immunoglobulin A, AST aspartate amino transferase, CPTN community periodontal index of treatment needs, ROS reactive oxygen species, IFN- γ interferon- γ , ALP alkaline phosphatase, PBI papilla bleeding index, HGF human gingival fibroblasts, LDH lactate dehydrogenase

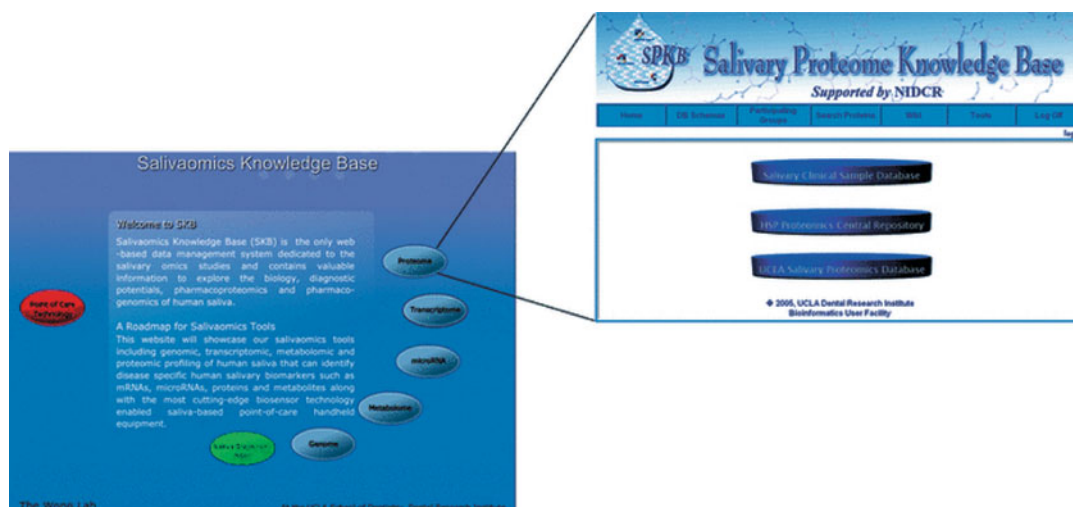


Fig. 12.45 The Salivaomics Knowledge Base and Salivary Proteome Knowledge Base from the Human Salivary Proteome Project supported by the National

Institute of Dental and Craniofacial Research (<http://www.skb.ucla.edu>) (Zhang et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

scriptomics, and metabolomics, which are now being widely implemented in studies of human disease. Furthermore, mining the data from multiple ‘omics’ studies can give deeper insight into biologic systems than can be obtained from any single ‘omics’ study, and thus, ‘omics’ databases such as PharmGKB (the Pharmacogenomics Knowledge Base, <http://www.pharmgkb.org>) and EPO-KB (the Empirical Proteomic Ontology Knowledge Base, <http://www.dbmi.pitt.edu/EPO-KB>) serve as important resources for emerging disciplines of system biology (Spielmann and Wong 2011).

Described as ‘Genomics meets Epidemiology’, the UK Biobank was created expressly to throw some light on the common diseases of middle and old age. More concretely, it aims to help disentangle the interactions between genomes and the lifestyles likely responsible for the epidemics of metabolic (e.g. diabetes type 2), genetic (e.g. cancer), and degenerative (e.g. Alzheimer’s disease) conditions. With half a million participants, aged 40–59 years, it is among the largest and—in terms of medical history, tissue collection, and archiving—the world’s most detailed survey of this particular population (Urdea et al. 2011).

Just recently, the **Salivaomics Knowledge Base (SKB)** has been created by aligning the salivary biomarker discovery and validation resources at UCLA (**Fig. 12.45**) with the ontology resources developed by the OBO (Open Biomedical Ontologies) Foundry (<http://www.obofoundry.org>), including the new Saliva Ontology (SALO). The SKB is a data repository, management system, and web resource constructed to support human salivary proteomics, transcriptomics, miRNA, metabolomics, and microbiome research. The SKB will provide the first web resource dedicated to salivary ‘omics’ studies. It will contain the data and information needed to explore the biology, diagnostic potentials, pharmacoproteomics, and pharmacogenomics of human saliva (SKB; <http://www.skb.ucla.edu>). Overall, the SKB will allow a systems approach to the utilization of salivary diagnostic technology for personalized medicine applications (Spielmann and Wong 2011; Ai et al. 2010, 2012).

SALO is a consensus-based controlled vocabulary of terms and relations dedicated to the salivaomics domain and to saliva-related diagnostics following the principles of the OBO Foundry. SALO was defined based on cross disciplinary

Fig. 12.46 Atomic Force Microscopy (AFM) images of saliva exosomes.

Exosomes (*panels b–f*) were adsorbed to WGA-coated mica surfaces.

(a) Topography images were obtained with the use of the Mac mode in water (negative control—no exosomes).

(b) A 3D AFM image of isolated exosomes adhering to a mica sheet. The bar denotes 200 μM .

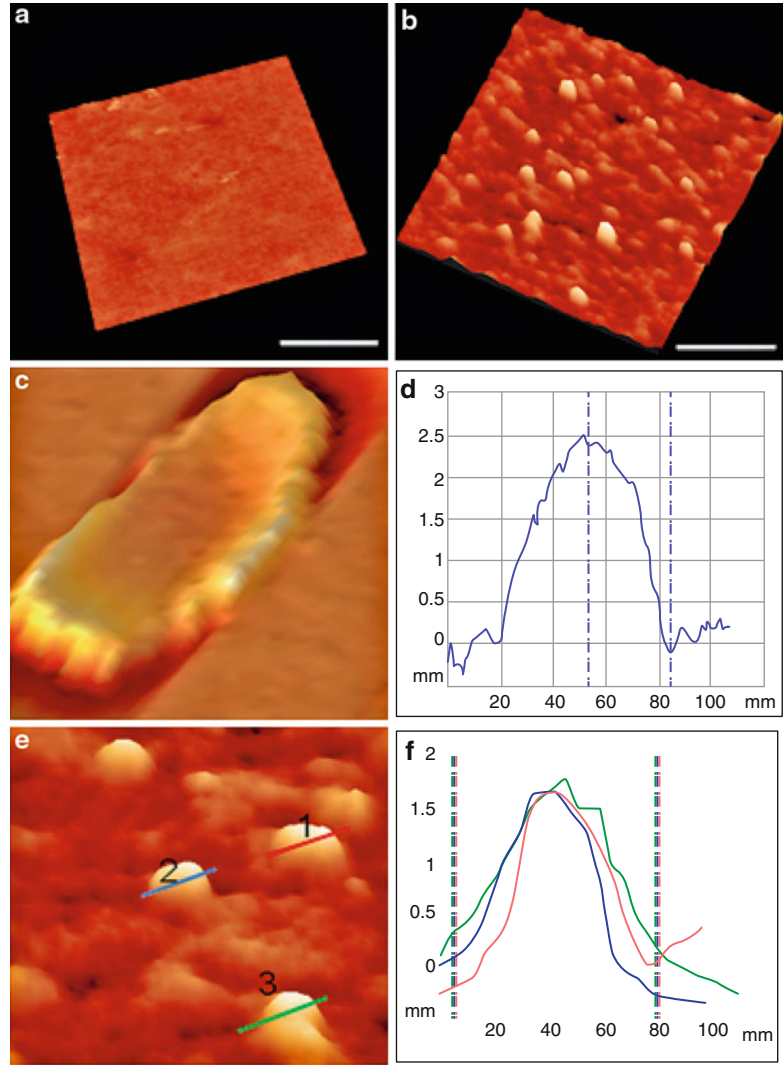
(c) A high-resolution single image of the exosome structure on the mica.

(d) Graphical representation of height and width of a single exosome.

(e) Size distribution of several saliva exosomes imaged with AFM.

(f) Graphical representation of the size distribution of exosomes showing near homogeneity with respect to height and width

(Chairoungdua et al. 2010, doi:10.1371/journal.pone.0008577.g002; Reprinted under the Creative Commons Attribution License)



interaction between saliva and protein experts, diagnosticians, and ontologists (SALO; <http://www.skbucla.edu/SALO/>). In order to improve development and testing of SALO, a corpus of saliva-relevant literature was incrementally developed in SKB to assist in identifying core terms, synonyms, and definitions for inclusion within the ontology and to provide examples of usage and links between SALO content and the corresponding items through their PubMed identifiers. A growing body of semantically enhanced web-enabled literature will be created within the SKB to support future research (Spielmann and Wong 2011).

A recent discovery is that salivary biomarkers (proteomic and transcriptomic) are contained in **exosomes**, which are small-membrane vesicles (Palanisamy et al. 2010; Sharma et al. 2010; Baum et al. 2011), 40–100 nm in diameter, corresponding to the internal vesicles present in multivesicular endosomes (MVEs), and have been known to be key for intercellular communication elsewhere in the immune system (Fig. 12.46) (Chairoungdua et al. 2010; Baum et al. 2011). The revelation that the salivary transcriptome and proteome are contained in exosomes not only explained their unusual stability in saliva but also provided a mecha-

nism whereby these nano-vesicles can mediate a signal transduction/communication axis to shuttle disease-specific exosomal contents (biomarkers, molecular constituents of mRNA, miRNA, and proto-oncogenic and pro-angiogenic proteins) to anatomical targets, including salivary glands. This systemic communicative capability of exosomes has been recently revealed, since they have been shown to be released from many different cell types to be important modulators of the immune system as well as oncogenesis. These recently recognized cellular communicative roles of exosomes in health and disease provide a compelling rationale for the hypothesis that systemic diseases shed exosomes containing disease-specific biomarkers that function as signalling nano-vesicles targeting different body sites, include including salivary glands and altered salivary biomarker profiles (Baum et al. 2011).

In addition, a catalogue (the **Human Oral Microbiome Database, HOMD**) of around 700 oral bacterial species identifiable by variations in 16S ribosomal RNA was published. This removed the problem posed by the impossibility of culturing many bacterial species to identify them. Allied to the high-throughput DNA sequencing methods now available, HOMD has facilitated the identification of bacterial biomarkers which might themselves be risk factors for diseases (Urdea et al. 2011).

12.5.2.5 Rapid Point-of-Care Diagnostics for Periodontal Disease

Periodontal surveillance and disease diagnosis will greatly advance over the coming years via the use of rapid point-of-care oral diagnostics. The drug-discovery process has been an excellent catalyst to link together novel therapeutics to emerging diagnostic disease biomarkers. This connection in biomedicine has led to the development of prototype rapid, accurate and ‘real-time’ assessment of multiple diseases, including periodontal disease. Novel technologies such as ‘lab-on-a-chip’ and microfluidic devices have the potential to manage complex oral fluids such as saliva and gingival crevicular fluid and to provide

a determination of a patient’s periodontal disease-risk profile, current disease ‘activity’, and response to therapeutic interventions. This approach should accelerate clinical decision-making and monitoring of episodic disease progression in a chronic infectious disease such as periodontitis (Giannobile 2007; Giannobile et al. 2009; Slavkin et al. 2011).

Single biomarker detection is not effective enough for accurate diagnosis and medical decisions because of the complexity of the human biologic system and the high possibility of false-positive and false-negative rates. The combination of multiple biomarkers could include highly discriminative nucleic acids, proteins, and small molecules such as metabolites (Spiellmann and Wong 2011).

The National Institute of Dental and Craniofacial Research has funded seven awards for the development of microfluidic and microelectromechanical systems for salivary diagnostics. Microfluidic and microelectromechanical systems are integrated arrangements comprising mechanical elements, sensors, actuators, and electronics on a common silicon substrate developed through microfabrication technology. These systems use small sample and reagent volumes coupled with integrated detection methods to perform analyses. The seven awards focused on the development of microfluidic and microelectromechanical system technologies to measure proteins, DNA, gene transcripts (mRNA), bacteria, electrolytes, and small molecules in saliva for point-of-care applications of human diseases. Researchers are designing ‘lab-on-a-chip’ prototypes (Zhang et al. 2009). The envisioned product is the ‘Oral Fluidic NanoSensor Test (OFNASET)’. The OFNASET is a handheld, automated, easy-to-use integrated system that will enable simultaneous and rapid detection of multiple salivary protein and nucleic acid targets (Fig. 12.47) (<http://www.skbucla.edu/>). This saliva biomarker detector can be used in dentist or health-care provider’s office for point-of-care disease screening and detection (Li et al. 2005; Lee et al. 2009). The intended use of the OFNASET is for the point-of-care multiplex detection of salivary biomarkers for oral cancer.



Fig. 12.47 UCLA's Oral Fluid NanoSensor Test (OFNASET). A point of care biosensor optimized for saliva detection of multiplex biomarkers, self-contained and being able to detect nucleic acids and proteins without

thermal cyclers and ELISA readers nor trained personnel (Lee et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

Gau and Wong (2007) have demonstrated that the combination of two salivary proteomic biomarkers (thioredoxin and IL-8) and four salivary mRNA biomarkers (SAT, ODZ, IL-8, and IL-1 β) can detect oral cancer with high specificity and sensitivity.

Barnfather et al. (2005) investigated the effect of immediate feedback from a point-of-care test for salivary nicotine metabolites in promoting smoking cessation and reduction in tobacco use. Participants completed a questionnaire on smoking, undertook a clinical examination, and received counselling in smoking cessation. Saliva samples were analysed at presentation and at 8 weeks for salivary nicotine metabolites using a

10-min semiquantitative point-of-care test. A higher smoking quit rate was achieved with the point-of-care test (23% cases vs. 7% controls; $P < 0.039$), and overall tobacco use also decreased (68% cases vs. 28% controls; $P < 0.001$). Incorporation of individualized personal feedback using a point-of-care test for salivary nicotine metabolites into a general practice-based smoking cessation programme increased quit rates by 17% at 8 weeks and reduced tobacco use (Barnfather et al. 2005).

Herr et al. (2007a) has developed a portable microfluidic device for detection of potential biomarkers of periodontal disease in saliva. The device performs rapid microfluidic chip-based

immunoassays (<3–10 min) with low sample volume requirements (10 μ l) and appreciable sensitivity (nM–pM). This microfluidic method facilitates hands-free saliva analysis by integrating sample pretreatment (filtering, enrichment, mixing) with electrophoretic immunoassays to quickly measure analyte concentrations in minimally pretreated saliva samples. Using 20 microl of saliva, Herr et al. (2007b) demonstrated rapid (<10 min) measurement of the collagen-cleaving enzyme matrix metalloproteinase-8 (MMP-8) in saliva from healthy and periodontally diseased subjects. The microfluidic chip has been integrated with miniaturized electronics; optical elements, such as diode lasers; fluid-handling components; and data acquisition software to develop a portable, self-contained device (Herr et al. 2007a).

Christodoulides et al. (2007) described a lab-on-a-chip (LOC) assay system, in which assays are performed on chemically sensitized beads populated into etched silicon wafers with embedded fluid-handling and optical detection capabilities. Using this LOC system, complex assays can be performed with small sample volumes, short analysis times, and markedly reduced reagent costs (Fig. 12.48). The use of LOC methodologies to assess the levels of interleukin-1beta (IL-1beta), C-reactive protein (CRP), and matrix metalloproteinase-8 (MMP-8) in whole saliva and the potential use of these biomarkers for diagnosing and categorizing the severity and extent of periodontitis were described. The detection limit of the LOC-based assay for CRP is significantly lower than that of ELISA, which is the current gold standard for high-sensitivity (hs) measurements of CRP. From a comparison of the two methods, it is clear that the LOC approach yields a five orders of magnitude lower limit of detection than that exhibited by the hsCRP ELISA method. Furthermore, the LOC assay was linear for three orders of magnitude, whereas ELISA was only linear for two orders of magnitude. Here, the LOC assay procedure demonstrates a detection limit at 5 fg/ml and a useful range between 10 fg/ml and 10 pg/ml. In contrast, the hsCRP ELISA method yields a detection limit

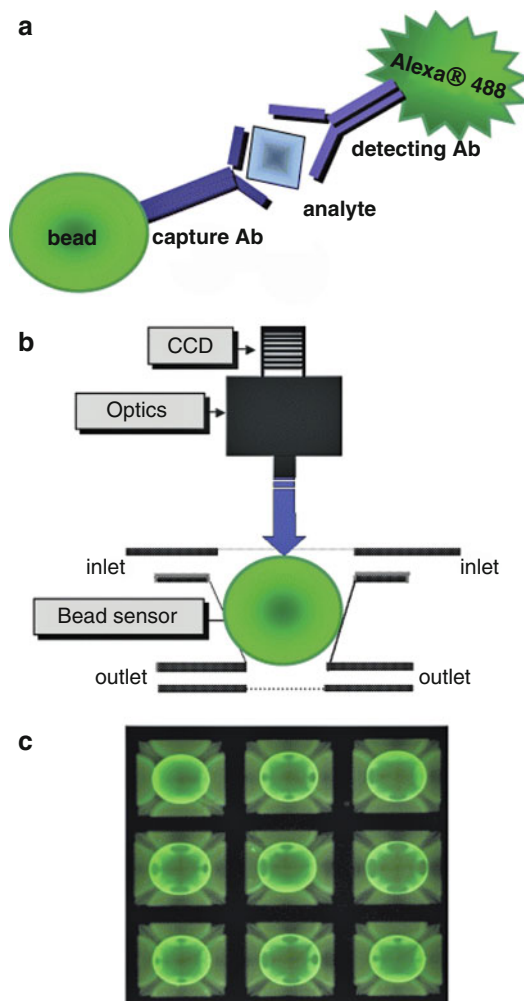


Fig. 12.48 The relevant immunocomplexes of the bead-based, sandwich-type of immunoassays on the LOC system are shown in (a). Here, an analyte-specific capturing antibody sequesters the antigen on the beads. Detection of the captured antigen is achieved in fluorescence mode with an Alexafluor-488@-conjugated detecting antibody. Completion of assays on the LOC system relies on the integrated function of its microfluidic and optical components (b). Here, during each assay, sample and detecting reagents are delivered to the beads via the top inlet of the flow cell, while the bottom drain provides an outlet for the direction of unspent reagents to a waste reservoir. The flow cell allows for the microscopic analysis and capture of signals generated on the array in conjunction with an optical station equipped with a CCD camera. Shown in (c) is a typical image of an array of beads captured in the last step of an LOC assay run (Christodoulides et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

of 2 ng/ml and a useful detection range between of 2 and 100 ng/ml CRP. Salivary levels of IL-1 β in the periodontal disease group were 2.6 times higher (as measured by LOC) and 1.5 times higher (as measured by ELISA) than the controls. Salivary MMP-8 levels were 2.0 times higher (as measured by LOC) and 2.6 times higher (as measured by ELISA) in the periodontal disease group compared with the healthy controls.

Multiplexing assay of biomarkers at the point of care is an elusive goal for molecular diagnostics. Recently, Wei et al. (2009) reported an electrochemical (EC) sensor for oral cancer detection based on the simultaneous detection of two salivary biomarkers: interleukin (IL)-8 mRNA and IL-8 protein. The EC sensor showed high sensitivity and specificity. The limit of detection is 3.9 fM for IL-8 mRNA and 7.4 pg/ml for IL-8 protein measured in saliva. With the multiplexing EC sensor, 56 clinical saliva samples were measured for oral cancer detection. Both IL-8 mRNA and IL-8 protein levels measured by the EC sensors show significant differences between the cancer sample and control sample. From the ROC analysis, the accuracies based on the EC sensor were both around 0.90 for the two biomarkers, which was very close to the results from the same batch of saliva samples measured separately by PCR/ELISA. Combined IL-8 mRNA and protein show better AUC compared with single biomarker. These results indicate that the EC sensor is not only an alternative detection method to PCR/ELISA. It provides fast, effective, and accurate multiplexing measurements for real clinical diagnostics (Wei et al. 2009).

Over the past decade, a US research team (Miller et al. 2010) has sustained efforts that combine and adapt the tools of nanomaterials and microelectronics for the practical implementation of miniaturized sensors, suitable for a variety of important applications. Here, two types of systems have been created. The first is based on a microbead array, wherein micro-pits within a silicon wafer are populated with a variety of chemically sensitized bead 'microreactors'. This sensor system is based on a bio-microelectromechanical systems platform and may be described as a 'chemical processing unit' in analogy to the cen-

tral processing unit that serves as the brains of a computer chip. Instead of handling electrical signals passing through conductors, as in the case of traditional circuits, the nano-biochip technology processes fluids so as to provide a digital fingerprint that can be correlated with the local chemical environment, detecting pH, electrolytes, metal cations, sugars, toxins, proteins, and antibodies. Building on this technology, a second class of miniaturized sensor system was developed that contains beads within etchings of stainless steel plates and utilizes a membrane capture element integrated into a fluidics structure (Miller et al. 2010 and references therein). The programmable bio-nanochip (PBNC) synergizes components and achievements from nanotechnology, clinical chemistry, bioinformatics, microfluidics, optics, image analysis, and pattern recognition to create a powerful new integrated measurement approach in a small device footprint. The PBNC system (Fig. 12.49) uses ~300- μ m-diameter bead sensors composed of agarose 'nanonets' that populate a microelectromechanical (MEMS) support structure with integrated microfluidic elements. The beads are an efficient and selective protein capture medium suitable for the analysis of complex fluid samples. This work uses microscopy and computational studies to probe the 3D interior of the beads. The PBNC features a flexible assay design and has a diverse collection of validated analytes, and its modular design allows for rapid deployment for new biomarker signatures. Assays for nucleic acids, proteins, and cells are arranged in the PBNC to create analytical test modalities specific to different disease types. Collectively, the modularity, flexibility, and ability to process and learn new biomarker signatures represent a biologic form of 'programmability'. Six major clinical trials are now active, involving the PBNC for major diseases in the areas of cardiac heart disease, oral cancer, ovarian cancer, and prostate cancer (Giannobile et al. 2011). Specifically, a multiplex platform designated 'cardiac arrest rapid diagnostic information using saliva' (CARDIUS), which uses four matched pairs of highly specific antibodies, each recognizing target antigens myoglobin (MYO), C-reactive protein (CRP), myeloperoxidase, and

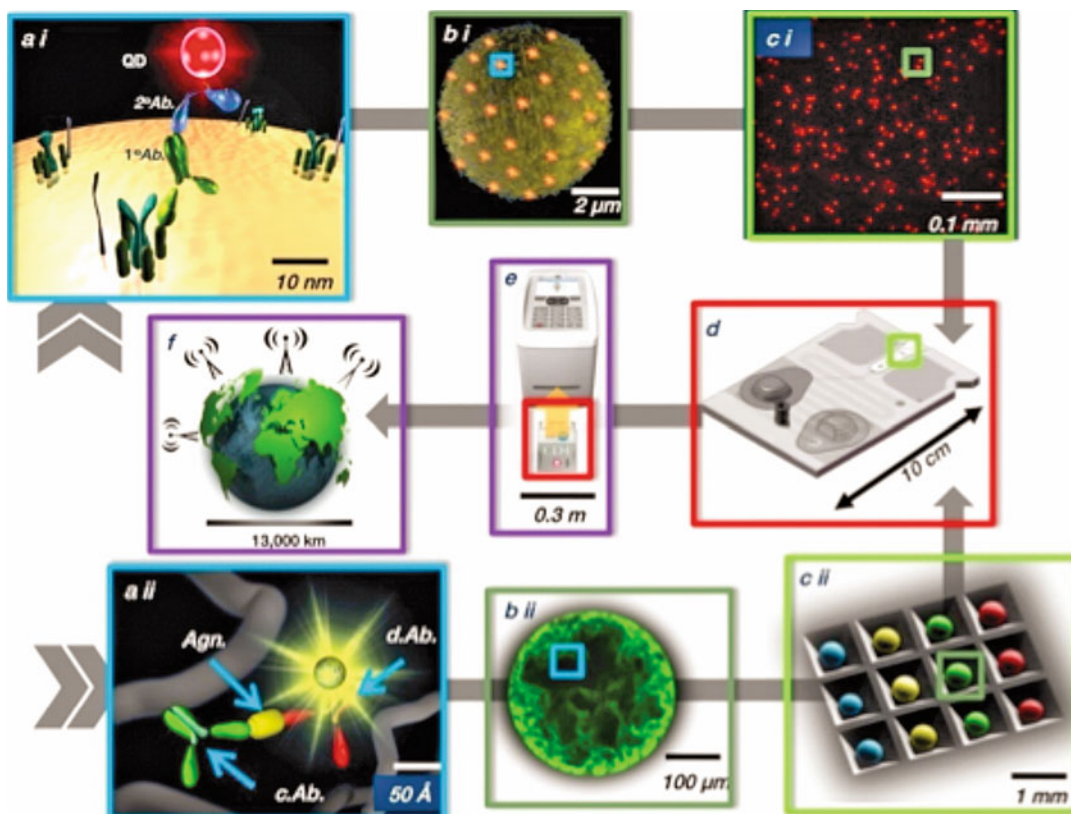


Fig. 12.49 Schematic illustration showing the main elements of the programmable bio-nano-chip system (PBNC). The universal detection system completes both cell-based

(top image sequence, **a i–c i**) as well as soluble analyte tests (bottom image sequence, **a ii–c ii**). (Giannobile et al. 2011. Reprinted with permission from Sage Publications)

IL-1 β , have demonstrated excellent acute myocardial infarction (AMI) screening capabilities (Floriano et al. 2009).

MyPerioID® PST® (OralDNA Labs® Inc.; www.oraldna.com): Designed to be fundamental elements of a patient's wellness plan, these salivary diagnostic tests measure periodontal disease infection and genetic risk factors for periodontal disease. MyPerioPath® offers genomic DNA testing that identifies the type and concentration of 13 specific periodontal pathogens found in a saliva sample. Saliva samples are mailed to the laboratory, and a detailed analysis is conducted, with a quantification of the specific pathogens and their risk factors. Additional clinical and medical risk factors are included as well as treatment considerations and reassessment recommendations. MyPerioID® PST® identifies

an individual's genetic susceptibility to periodontal disease by testing for the presence of two interleukin-1 polymorphisms. Both tests can be used as part of the baseline data for any patient at increased risk for periodontal infections as well as for patients who are unresponsive to their current treatment (Draper 2010).

As Giannobile et al. (2011) revealed, emerging clinical applications of **lab-on-a-chip (LOC)** technologies as **point-of-care (POC)** diagnostics developed for systemic diseases are now being readily applied to periodontology (Christodoulides et al. 2007; Herr et al. 2007a, b; Miller et al. 2010). The use of proteomic or multi-analyte approaches for the identification of periodontal diseases offers significant potential for providing periodontal disease 'signatures' for risk (Ramseier et al. 2009; Gonçalves Lda et al. 2010; Haigh et al. 2010; Giannobile et al. 2011). The use of

optimized point-of-care devices for periodontal surveillance will probably require less training and fewer resources than current diagnostic tests, could lead to better utilization of skilled clinicians for simpler and less intensive treatment, and may result in a more cost-effective health-care delivery. The portable, easy-to-use diagnostic tools will allow patients to be screened for periodontal disease in settings other than the dental practice, such as at a physician's office or in the home, allowing patients to be screened for early disease detection and directed for possible treatment. Periodontal oral diagnostic devices will also enable screening of large populations (especially the underserved communities and resource-poor areas) around the globe more quickly and effectively than the current inefficient screening approaches. The potential to sample various populations will help to identify at-risk groups more effectively and increase access to treatment for those most at need, improving public health (Giannobile et al. 2009).

12.5.3 Source for Samples of the Different Diagnostic Tests: GCF Components

Gingival crevice fluid (GCF) is a complex mixture of substances derived from serum, leukocytes, structural cells of the periodontium, and oral bacteria (Fig. 12.50). These substances possess a great potential for serving as indicators of periodontal disease and healing after therapy. Some of the suspected periodontopathogens, such as *Porphyromonas gingivalis* and *Treponema denticola*, produce broad-spectrum neutral proteinases as part of their virulence arsenal. These proteinases may be detected in plaque and GCF samples of periodontitis patients (Uitto 2003).

12.5.3.1 Methods of Collection and Measurement of the Volume of Gingival Crevicular Fluid (GCF)

Several techniques have been employed for the collection of GCF, and the technique chosen will depend upon the objectives of the study as each

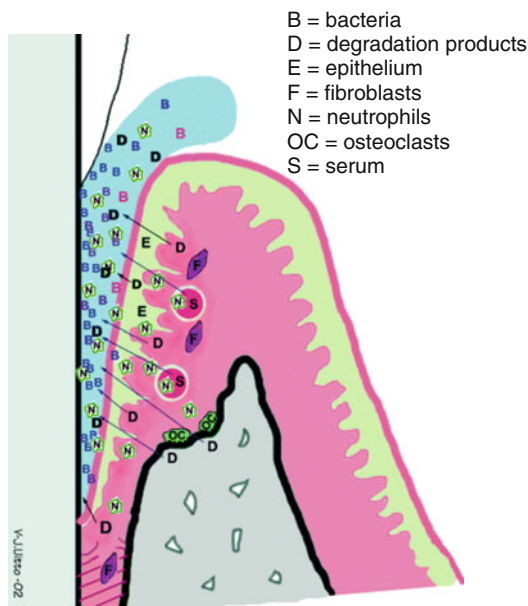


Fig. 12.50 Gingival crevice fluid – a window to periodontal disease. Gingival crevice fluid is composed of substances derived from serum, leukocytes, bacteria, activated epithelial cells, connective tissue cells, and bone cells. Tissue destruction during periodontal inflammation results in production of tissue fragments and growth factors released from tissue stores. All these substances reflect the periodontal disease process and can be potentially used as indicators of periodontal condition (Uitto 2003). Reprinted with permission from John Wiley & Sons, Inc.)

technique has advantages and disadvantages. The techniques can be divided into three basic strategies, subject to various modifications in their application by different authors (Griffiths 2003).

- **Gingival washing methods:** In this technique, the gingival crevice is perfused with an isotonic solution, such as Hanks' balanced salt solution, usually of fixed volume. The fluid collected then represents a dilution of crevicular fluid and contains both cells and soluble constituents such as plasma proteins. It is a technique that has usually only been applied to the maxillary arch, presumably because of the difficulties of producing a technically satisfactory appliance for the mandibular arch. The major disadvantage of this technique is that all fluid may not be recovered during the aspiration and re-aspiration procedure. Thus accurate quantification of GCF volume or composition is

not possible as the precise dilution factor cannot be determined (Griffiths 2003).

- **Capillary tubing or micropipettes:** Small capillary tubes can be placed within or at the orifice of the gingival crevice. By capillary action, fluid enters the tube. The volume of fluid entering the tube can be precisely measured, but it was noted that there is difficulty in collecting gingival crevicular fluid if the fluid is viscous and due to the prolonged harvesting time (collection times from an individual site may exceed 30 min), resulting in considerable trauma to the gingival blood vessels, changing the gingival crevicular fluid flow and the compositions of its components (Gustafsson 1996). A further complication of this technique is the difficulty of removing the complete sample from the tubing. This has either been forced out with a jet of air or by passing a larger fixed volume of a diluting solution through the capillary or, more usually, by centrifugation of the tube (Griffiths 2003).
- **Absorbent filter paper strips:** The most clinically applicable method uses precut methylcellulose filter paper strips that are placed in the sulcus (Fig. 12.51). The fluid absorbed on the strips is then eluted and analysed. This approach offers the potential of a noninvasive means of assessing the host response in periodontal disease. The filter strip method of fluid collection can, however, be time consuming and technique sensitive. Attention must be given to strip placement so that the sample is relatively free of plaque (supragingival plaque is removed) and there is no contamination by saliva and blood. Also, the strip must remain in the sulcus long enough to obtain an adequate sample of fluid. Concern has also been raised about the approach to reporting data that examined the levels of host mediators in GCF. The problem is related to the available volume of GCF. Unlike the analysis of serum, where the sample of fluid is a small part of the total fluid volume, sampling of GCF often collects the entire volume of fluid at the sampled site, and this volume varies from tooth site to tooth site. As a result, Lamster et al. (1988) and Lamster (1997) developed an approach to



Fig. 12.51 Gingival crevicular fluid is most often collected with methylcellulose filter paper strips inserted into the crevice. Standardization is obtained with timed sampling (30 s) (Lamster and Ahlo 2007. Reprinted with permission from John Wiley & Sons, Inc.)

GCF sampling that standardizes the time of collection and reports the data as total amount (or activity) in the timed sample (Lamster and Ahlo 2007).

Different techniques were described for sampling the contents of the periodontal pocket (Griffiths 2003). The sampling of subgingival bacteria seems to be suitable with curettes or paper points (Jervøe-Storm et al. 2007), whereas cytokines and host enzymes were usually collected with filter paper strips. Recently, Guentsch et al. (2011) has investigated a method that could be used for different purposes (e.g. microbiota and immunologic variables). It was demonstrated that the washing technique is an alternative sampling method of GCF for special purposes when sampling by paper-based methods fails. Paper points are suitable for the determination of the microflora and are recommended for daily microbiologic analysis in dental practice. Paper strips are the method of choice for most biomarkers in immunologic studies; a combined determination of periodontopathic bacteria seems possible (Guentsch et al. 2011).

Methods of Collection

The methods of collection may be broadly divided into the intracrevicular and the extracrevicular techniques. The former depends on the strip being inserted into the gingival crevice, whereas in the latter the strips are overlaid on the gingival crev-

ice region in an attempt to minimize trauma. The intracrevicular method is the method used most frequently and can be further subdivided depending upon whether the strip is inserted just at the entrance of the crevice or periodontal pocket or whether the strip is inserted to the base of the pocket or 'until minimum resistance is felt'. In shallow pockets or healthy crevices, they probably represent the same thing (Griffiths 2003).

Methods of Estimating the Volume Collected

In many early studies, the only purpose of the investigation was to determine the amount of GCF produced at a given site. This relegated the sampling of GCF to an additional clinical measurement. The amount of GCF collected on a strip was assessed by the distance the fluid had migrated up the strip. Further accuracy was achieved by staining the strips with ninhydrin or fluorescein to mark the area where GCF had accumulated. The staining techniques have a number of disadvantages; firstly, they are not easily applied at the chairside. The inevitable delay in measuring the strip may result in increased variation in the reported volume as a result of evaporation. Secondly, the staining of the strips for protein labelling prevents further laboratory investigations of the components of GCF, effectively limiting the technique to that of volume determination. The introduction of an electronic measuring device, the Periotron®, has allowed accurate determination of the GCF volume and subsequent laboratory investigation of the sample composition. The instrument measures the effect on the electrical current flow of the wetted paper strips. It has two metal 'jaws' which act as the plates of an electrical condenser. If a dry strip is placed between the 'jaws', the capacitance is translated via the electrical circuitry and registers 'zero' on the digital readout. A wet strip will increase the capacitance in proportion to the volume of fluid, and this can be measured as an increased value in the readout. The technique is rapid and has no discernible effect upon the GCF sample (Griffiths 2003).

From most of the studies on calibration of the Periotron®, the following recommendations can be made (Griffiths 2003):

- Each machine needs its own calibration, as machines differ markedly in their range.
- Serum is a suitable material to generate a calibration curve.
- An accurate syringe and some form of standardization of dispensing duplicate volumes are essential.
- Duplicate volumes are required, but it is not necessary to add additional replicates.
- A full range of volumes from 0.1 to 1.2 μl is necessary to generate an accurate curve.

Regarding the appropriate number of repetitions for the generation of a reliable calibration curve, suggestions and clinical applications seem to vary. Preshaw et al. (1996) recommended that each calibration volume should be repeated at least in triplicates for Periotron 6000®, whereas Deinzer et al. (2000) suggested repetition of five times. However, Griffiths (2003) suggested that duplicate volumes were sufficient for Periotron 8000®. Tözüm et al. (2004) showed that, although statistical analysis was not performed, the mean, minimum, and maximum PU values for additional 5× or 20× replicates did not seem to be far different than the values of 3 ×, and additional replicates did not seem to add further benefit. Thus, their findings may support the reliability of triplicate readings for Periotron 8000®, as similar to Periotron 6000® (Preshaw et al. 1996).

An alternative approach, involving the weighing of strips before and after sample collection, has been adopted. This has been successful but requires a very sensitive balance to estimate the very small amounts of fluid which may be collected from a healthy crevice. As with the protein staining and assessment of wetted area techniques described earlier, evaporative losses because of delays in determining the volume may distort the volumes obtained. Indeed, if a filter paper strip with either GCF or a known volume of serum is placed on a balance without being included within a sealed container, evaporation can be observed by following the decrease in weight recorded on the balance (Griffiths 2003).

As has been illustrated, many sampling protocols that provide an estimate of GCF flow have

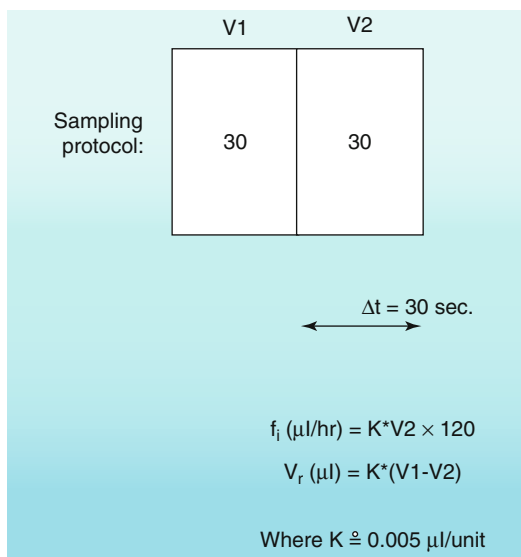


Fig. 12.52 Proposed simplified protocol to provide a chairside measure of gingival crevice fluid (GCF) flow (f_i) and resting volume (V_r). The sampling protocol requires taking two sequential 30-s GCF samples (Goodson 2003. Reprinted with permission from John Wiley & Sons, Inc.)

been tested. Goodson (2003) suggested that a simple sampling protocol for routine clinical use measures the volume in Periotron units of two 30-s GCF samples (Fig. 12.52). The volume of the second sample (V_2) in Periotron units multiplied by 0.6 (the calibration factor, approximately $0.005 \mu\text{l/unit}$, multiplied by $60.60/30$) is the flow rate in $\mu\text{l/h}$. The first volume (V_1) minus the second volume (V_2) multiplied by 0.005 (the calibration factor) is the resting volume V_r . Given the evidence that change in GCF flow may be a sensitive measure of local inflammation, this type of evaluation could increase the diagnostic potential of this chairside measure. Because of the time required for a single measurement, it is unlikely that this would compete with probe measurements for full-mouth evaluation. Nevertheless, when it becomes necessary to focus on treatment of an individual problem tooth, GCF flow measurement could provide added benefit in establishing a diagnosis and monitoring the response to therapy (Goodson 2003).

Problems with GCF Collection and Data Interpretation (Griffiths 2003)

1. **Contamination:** The major sources of contamination of GCF samples would be blood, saliva, or plaque. Frank blood contamination is usually dealt with by discarding the sample and removing the data from analysis.
2. **Sampling time:** The early literature from the Periotron® suggested that filter paper strips should be left in place for 5 s. Alternative approaches to sampling techniques have been developed in an effort to increase the volume of GCF sample available for subsequent laboratory analysis.
3. **Volume determination:** As highlighted earlier, evaporation is considered to be a significant problem in accurate volume determination of GCF samples. This is particularly the case as the total volumes collected are usually less than $1 \mu\text{l}$ and more often than not are less than $0.5 \mu\text{l}$. Concerns are also expressed regarding the accuracy of the calibration graph produced for the Periotron®, particularly with respect to small volumes.
4. **Recovery from strips:** Having collected the GCF sample and determined its volume, the samples are usually then required for some investigation of the composition of GCF. To achieve this, it is necessary to recover the GCF from the filter paper strips, and initial work indicated that protein recovery was close to 100% using a centrifugal elution technique. A variety of other methods of elution have been employed, but it is essential in all instances to determine the percentage recovery from the original samples.
5. **Data reporting:** In view of the extensive laboratory studies, in a wide variety of fields, which report significant dose–response effects, it would seem sensible to report GCF data as both total amount and concentration. The inclusion of the volume data would then allow the reader a clear view of what is occurring. Furthermore, some estimation of the biological effect of the amount or concentration of the constituent examined would appear to be desirable.

12.5.3.2 Determinants of the Volumetric Features of the GCF

Clinical periodontal status is among the major determinants of the volumetric features of GCF. Studies have shown that probing depth (PD), presence/severity of gingival inflammation, and periodontitis affect the biodynamics of GCF (Hatipoğlu et al. 2007; Ozkavaf et al. 2000). In reviewing the currently available data in which GCF flow measurements can be inferred, Goodson (2003) showed that there is clearly a disease-related spectrum of values. Shallow sulci in healthy subjects have GCF flow rates of 3–8 $\mu\text{l/h}$. Pockets with intermediate periodontal disease have GCF flow rates of approximately 20 $\mu\text{l/h}$. GCF flow at sites with advanced periodontal disease as high as 137 $\mu\text{l/h}$ has been observed. The evidence from experimental gingivitis studies that GCF flow can increase dramatically with no change in resting volume suggests that this may be a sensitive early indicator of gingival inflammation. Evidence from treatment studies indicates that GCF flow can also be used to evaluate treatment response. The currently available data in which GCF resting volume values can be inferred indicate that this too has a disease-related spectrum of values. Shallow sulci in healthy subjects have resting volumes in the order of 0.06 μl . Pockets with periodontal disease have resting volumes from 0.4 to 1.5 μl . This agrees with values estimated by radioactive isotope dilution to be in the range of 0.24–1.56 μl (Goodson 2003).

Obviously, GCF volume differs among sites, and fluid cannot be sampled in the entire mouth. Thus, identification of the appropriate subsets of sites 1 and **appropriate selection of the sampling site** become critical volumetric determinants of a sound methodology. Analysis of GCF-related studies has revealed that selection of the sampling site is based on various different criteria (e.g. PD, clinical periodontal status, susceptibility/resistance to periodontal destruction, anatomical relation of sites, and distinct teeth groups). It can also be seen that the preference of different sampling sites such as maxillary, mandibular, posterior, anterior, or various index teeth and approximal, labial–buccal, or palatal–lingual approaches are

all available. Appropriate location of the sampling area is vital because sampling site has a consistent effect on both resting volume (V_r) and GCF flow. The clear site-specific V_r variations are attributed to the differences in the surface area of the crevices, contact area of the strips with the sulcus/pocket epithelium, and unique dimensional features of distinct teeth groups. Furthermore, gravity, accessibility of the site, and extent of risk of contamination with plaque and saliva are also of specific concern (Hatipoğlu et al. 2007 and references therein). When the analysis of the possible impact of the clinical periodontal status and the distinct location of sampling sites on fluid volume was performed, it was revealed that although volume increased in a disease-related pattern (healthy < gingivitis < periodontitis; $P < 0.05$), the distribution range of volume was widespread, with prominent overlaps between the different clinical periodontal conditions. Multirrooted teeth presented more fluid volume, and even mesio-buccal or disto-buccal sites exhibited some volumetric differences ($P < 0.05$). Constant correlations between volume and clinical parameters could be observed only at gingivitis sites ($P < 0.05$) (Hatipoğlu et al. 2007).

Gingival fluid production is increased by mastication of course foods, tooth brushing, gingival massage, ovulation, hormonal contraceptives, and **smoking** (Nisengaard et al. 2006). It has been reported that the GCF volume was lower in smokers than in nonsmokers (Hedin et al. 1981; Persson et al. 1999; Gomes et al. 2009). The GCF volume was statistically significantly increased compared to the baseline at 5 days after smoking cessation ($P < 0.01$). However, the GCF volume of smokers until 2 weeks after smoking cessation was statistically significantly low compared to nonsmokers ($P < 0.01$) (Morozumi et al. 2004).

Volumetric measures are also highly sensitive to **sampling techniques** (Lamster et al. 1986; Persson and Page 1990; Shapiro et al. 1979; Ozkavaf et al. 2001). Because different sampling techniques, even the placement of paper strips with different extensions, and the sampling time have been shown to alter actual volume/flow of GCF to various magnitudes, methodological considerations mainly emphasize the importance of

the clinical sampling technique and sampling time (Tözüm et al. 2004 and references therein). It has been well shown that mechanical injury during sampling can increase vascular and sulcular epithelium permeability and the flow rate of the fluid, thus diluting the actual GCF (Lamster et al. 1986; Persson and Page 1990; Shapiro et al. 1979; Ozkavaf et al. 2001; Hatipoglu et al. 2007).

However, if not followed by precise volume quantification, a successful clinical GCF sampling on its own may not ensure a reliable outcome due to additional factors that operate following *in vivo* GCF sampling and that have the potential to interfere with the accuracy of the process of volume quantification. Fluid evaporation, room temperature, and humidity may interfere with the process of calibration of the device and the maintenance of the reliability of the calibration curve, which are also crucial issues for accurate volume quantification (Tözüm et al. 2004 and references therein).

The site-specific nature of GCF and the great volumetric fluctuations among sites emphasize the need for further attempts to standardize the GCF sampling methodology.

12.5.3.3 Reliability of GCF Volume Determination

Surprisingly few authors focused on this problem and analysed the reliability of GCF volume determination. In a group of 18 subjects, Stewart et al. (1993) explored gingival crevicular fluid (GCF) volumes as measured by the Periotron. Collections of GCF from both the buccal and lingual surfaces of 6 teeth from 18 subjects with good to moderate plaque accumulation were accomplished. At a later time, these same surfaces were retested for a comparison with the initial value for reliability determination. When the GCF values from the 12 tooth surfaces were averaged for each subject, differences between the two measurements improved markedly. Sixty-one percent of these subject pairs differed by less than 20%. As these reliability coefficients, however, refer to averaged measures, they may overestimate the reliability of GCFV determination at single sites: Aggregating data by averaging increases reliability (Steyer and Schmitt

1990). Therefore Deinzer et al. (2000) assessed the retest-reliability of GCF volumes measurements at single sites. For assessment of the retest-reliability, GCF volume was determined twice with an intertrial interval of 5 min. The retest-reliabilities were $r_{tt}=0.85$ for tooth 11 disto-buccal and $r_{tt}=0.84$ for tooth 26 mesio-buccal (Deinzer et al. 2000).

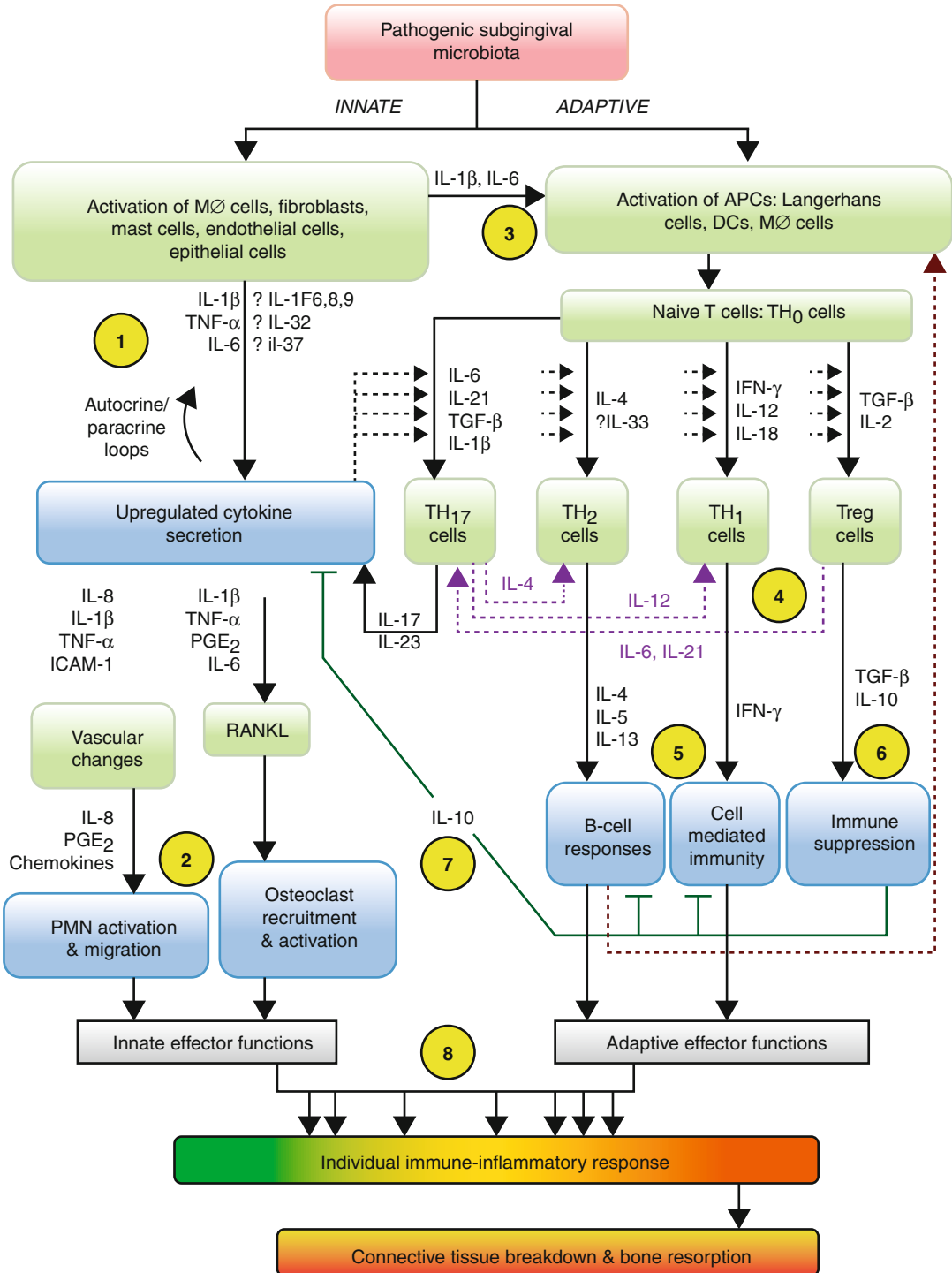
12.5.3.4 Gingival Crevicular Fluid and Periodontal Disease

The host-derived substances in GCF include antibodies, cytokines, enzymes, and tissue degradation products. The antibodies in gingival crevicular fluid are comprised of both locally and systemically synthesized molecules, and they reflect periodontal colonization by particular microbial species. Antibodies are key players in the protection of the body against pathogenic bacteria. However, they may also be crucial in eliciting destructive inflammatory reactions in the periodontium. Bone-specific markers such as the collagen telopeptide fragments and osteocalcin in GCF may reflect periodontal bone resorption. Proteolytic enzymes are released in periodontal tissues during inflammation from leukocytes, activated structural cells of epithelium, and connective tissue. Of these enzymes, matrix metalloproteinases (MMPs) have the main responsibility for degradation of periodontal tissues. Recent studies have shown, however, that MMPs also have specific functions in controlling tissue defence reactions and repair. In an attempt to find a specific MMP indicator for periodontal diseases, focus should be directed to proteinases that are released during the active tissue destruction episodes of the diseases rather than those released from neutrophils throughout the inflammatory process. Also active rather than total (latent plus active) enzyme activities should be measured (Uitto 2003).

Numerous cytokines are present that play key roles in the pathophysiology of periodontitis. The strongest evidence for cytokines functioning in networks exists for interleukin (IL)-1 β , tumour necrosis factor- α (TNF- α), IL-6, and RANKL. More recent data highlight a pro-inflammatory role for IL-10 in periodontitis, in addition to its

anti-inflammatory function. Cytokine networks are likely to control T-cell development and plasticity in periodontal lesions. Variations in these

networks may have a major impact on T-cell phenotype and disease expression (**Fig. 12.53**) and hence warrant further investigations (Ford et al.



2010; Kinane et al. 2011). Moreover, it has been shown that different subgingival biofilm profiles are associated with distinct patterns of GCF cytokine expression in both aggressive and chronic periodontitis subjects (Teles et al. 2010a, b).

On the basis of our current understanding of the complexity of periodontitis, the identification of one single diagnostic marker for all forms of periodontal disease seems illusory. Nevertheless, researchers have been searching actively for unequivocal markers of periodontitis in gingival crevice fluid to develop a simple test, to be used chairside, and to determine whether a patient suffers from periodontitis and needs therapy, as opposed to another patient who needs no

intervention even though he has gingivitis (Loos and Tjoa 2005).

A substantial number of studies throughout the 1980s and the 1990s explored the predictive ability of GCF components for identification of progressive periodontal lesions. While individual GCF components produced positive predictive values that were superior to individual clinical measures, these studies focused largely on the prediction of periodontitis at the site level rather than the identification of high-risk groups and individuals. Data from the selected studies investigating potential biomarkers in GCF in relation to periodontal disease are outlined in **Table 12.9** (Buduneli and Kinane 2011). In their review, Loos and Tjoa (2005) have searched the literature

Fig. 12.53 Cytokine networks in periodontal diseases. Schematic to illustrate the multiple interactions between cytokines and cellular functions in periodontal diseases. (1) Resident and infiltrating cells in the periodontium respond to MAMPs signalling via PRRs by production of cytokines as an early step in innate immune responses. Cytokine upregulation is sustained by autocrine and paracrine feedback loops. (*Note: question marks (?) indicate more speculative suggestions about the role of specific cytokines in periodontal pathogenesis as known at present.*) (2) Upregulated cytokine activity leads to vascular changes, PMN activation and migration, and ultimately, osteoclastogenesis and osteoclast activation. (3) Cytokines produced in innate responses contribute to activation of APCs. These present specific antigens to naive CD4⁺ T-cells (Th0 cells), which differentiate into CD4⁺ effector T-cells (e.g. Th1, Th2, Th17, Treg cells) according to the local cytokine milieu (as indicated by the groups of four parallel horizontal grey dotted arrows). For example, Th0 cells differentiate to Th17 cells under the influence of IL-6, IL-21, TGF- β , IL-1 β . (APCs are also activated by B-cells, which are themselves activated at a later stage in the cytokine network – indicated by brown dotted arrow at right edge of figure – an example of the complexities of sequential feedback loops that develop.) (4) Th1 and Th2 cells have a relatively stable phenotype, but other T-cell subsets can exhibit functional plasticity under the influence of different cytokine environments (indicated by purple dotted arrows). For example, Th17 cells can develop into Th1 cells under the influence of IL-12, and into Th2 cells under the influence of IL-4. (5) Different T-cell subsets are associated with various cytokine secretion profiles which regulate different aspects of immune responses and contribute to upregulated cytokine activity. For example, Th1 cells secrete IFN- γ (which activates cell-mediated immunity), and Th2 cells regulate antibody-mediated (humoral) immunity

through secretion of cytokines IL-4, IL-5 and IL-13. Cytokines produced by different T-cell subsets increase their further secretion in positive feedback loops, and also inhibit development of other T-cell subsets (e.g. IL-4 from Th2 cells inhibits Th1 development, and IFN- γ from Th1 cells inhibits Th2 T-cell subsets). (6) Treg cells secrete TGF- β and IL-10 which have immune suppressive functions. For example, IL-10 suppresses Th1 and Th2 responses, suppresses M ϕ and DC function, and downregulates cytokine production in various cells (Th1 cells, Th2 cells, PMNs, NK cells. (Suppressive effects are indicated by flat-ended green lines.) (7) IL-10 functions as a regulatory mediator, but can also exhibit other activities such as activation of B-cells. The different aspects of IL-10 biology (immunosuppressive versus immunostimulatory) likely depend on the local cytokine environment. These dual roles of IL-10 are indicated by the green (inhibitory) line and the black (stimulatory) arrow. (8) The sum total of innate and adaptive effector functions results in an immune-inflammatory response, the precise nature of which will vary from person to person (as indicated by the multiple grey arrows, with some patients being more susceptible to disease than others), and over time within an individual. In this case, the black arrow indicates an individual who has a pro-inflammatory response that leads to connective tissue breakdown and bone resorption. APC antigen presenting cell, DC dendritic cell, IFN- γ interferon- γ , ICAM-1 intercellular adhesion molecule-1, IL interleukin, MAMP microbe-associated molecular pattern, M ϕ cell of monocyte/macrophage lineage, NK natural killer, PGE₂ prostaglandin E₂, PMN polymorphonuclear leucocyte, PRR pattern recognition receptor, RANKL receptor activator of NF- κ B ligand, TGF- β transforming growth factor- β , Th T-helper cell, TNF tumour necrosis factor, Treg T regulatory cell (Kinane et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

Table 12.9 Studies evaluating possible biomarkers in gingival crevicular fluid (GCF) samples (Buduneli and Kinane 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Akalin et al. (2005)	Cross-sectional	26 CP, 18 controls		Routine clinical parameters	GCF, biopsy	SOD	Increased in gingiva, no difference in GCF
Akalin et al. (2007)	Cross-sectional	36 CP, 28 healthy			Serum, saliva, GCF	Lipid peroxidation, malondialdehyde, total antioxidant status	LPO and TOS may play an important role in periodontitis
Baltacıoğlu et al. (2008)	Cross-sectional	33 CP, 24 healthy			Serum, GCF	Total protein, protein carbonylation	Elevated protein carbonylation may be a sign of oxidative stress in CP
Buduneli et al. (2005)	Cross-sectional	20 gingivitis, 20 CP, 20 healthy controls.			GCF, serum	t-PA, u-PA, PAI-1, PAI-2	GCF PAI-2 concentrations were higher in CP, gingivitis than healthy controls. Periodontal disease rather than smoking seems to affect PA system
Chapple et al. (2007)	Prospective	18 CP	3 months after SRP		GCF, plasma	TAOC	GCF TAOC was lower in CP, increased after SRP
Chen et al. (2010)	Cross-sectional	25 CP, 24 healthy			GCF, serum	PAF	Higher PAF in GCF, serum in CP
Dezerega et al. (2010)		15 progression, 18 CP, 10 healthy			GCF	MCP-3	Higher in CP, higher in active sites than inactive sites, may be a diagnostic marker
Dutzan et al. (2009)	Cross-sectional	106 moderate to advanced CP			GCF	IFN- γ	Higher in active sites, may be an indicator for progression
Emingil et al. (2006b)	Cross-sectional	18 GAgP, 29 CP, 20 gingivitis, 20 healthy controls.		PD, CAL, BOP, PI	GCF	Laminin-5 gamma2-chain	CP had elevated laminin-5 gamma2-chain levels
Fitzsimmons et al. (2010)	Cross-sectional, population based	430 periodontitis, 509 healthy controls			GCF	IL-1 β , CRP	Higher levels in periodontitis, GCF levels of IL-1 β , CRP may indicate higher susceptibility

Gapski et al. (2009)	Multi-centre, prospective,	CP SDD + surgery Placebo + surgery	12 months	1° outcome: CAL 2°	GCF	ICTP	SDD decreased ICTP effectively Provided better
Garg et al. (2009)	Intervention	20 healthy 20 gingivitis 20 CP (also after SRP)	6–8 weeks after SRP	GI, PD, CAL	GCF	Cathepsin K	Cathepsin K increased in CP; decreased with SRP
Golub et al. (2008)	Prospective, controlled	CP; Postmenopausal women SDD: 64 Placebo: 64	2 years		GCF	ICTP, MMP-8, MMP-1, MMP-13, IL-1 β	SDD decreased collagenase activity, ICTP
Grant et al. (2010)	Intervention	20 healthy, 20 CP	Before and 3 months after SRP		GCF	Reduced glutathione, oxidized glutathione	GSH, GSSG lower in CP GSH: GSSG ratio increased after SRP
Holzhausen et al. (2010)	Cross-sectional, and intervention	40 moderate CP 40 severe CP 40 healthy controls	At baseline and also 1 month after SRP		GCF	Trypsin-like activity, TNF- α IL-1 α IL-6, IL-8, PAR ₂	PAR ₂ was higher in CP than controls. IL-1 α , -6, -8 TNF- α , total proteolytic activity were all higher in CP
Ikezawa-Suzuki et al. (2008)	Intervention	35 CP	At baseline and after SRP		GCF	TNF- α , TNF receptor types 1 and 2	Ratio of TNF R2: R1 increased after treatment, may be related to clinical periodontal data
Jepsen et al. (2003)	Cross-sectional	54 teeth with CP, 11 experimental gingivitis			GCF	lysylpyridinolone, hydroxylysylpyridinolone	Increased hydroxylysylpyridinolone, Lysylpyridinolone may be indicators of ongoing destruction
Kardeşler et al. (2008)	Cross-sectional	17 DM + periodontal disease, 17 healthy + periodontal disease, 17 healthy controls			GCF	PGE ₂ , IL-1 β , t-PA, PAL-2	PGE ₂ , t-PA were higher in DM, and periodontal disease groups than in healthy controls
Kurtiş et al. (1999)		24 CP, 24 healthy		PD, PI, GI, BOP, CAL	GCF	IL-6	Higher in CP, no correlation with clinical data
Lamster et al. (1994)					GCF, saliva, serum	β -glucuronidase	GCF levels correlated with CAL

(continued)

Table 12.9 (continued)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Lin et al. (2005)	Cross-sectional	14 patients (62 sites divided into 4 groups by PD, BOP).		PD, BOP, CAL	GCF	Oncostatin M, IL-6	Total amounts but not concentrations correlated with disease severity.
Mogi and Ootogoto (2007)	Cross-sectional	66 CP; 20 mild, 24 moderate, 22 severe, 19 healthy controls			GCF	Cathepsin K	Increased in periodontitis versus control subjects, (+) correlation between cathepsin K and RANKL suggesting that both contribute to osteoclastic bone destruction
Mogi et al. (2004)	Cross-sectional				GCF	Cathepsin K, RANKL	Both increased in CP
Nakamura-Minami et al. (2003)	Intervention	21 CP	4 weeks after SRP	Routine clinical parameters	GCF	α -1 protease inhibitor, secretory leukocyte protease inhibitor	Both seem to be related with healing after SRP
Oringer et al. (2002)	Intervention	48 CP; SRP+placebo, SRP+minocycline	Before, 1,3,6 months after SRP		GCF	ICTP, IL-1	ICTP, IL-1 correlated with clinical data, may be disease markers
Palys et al. 1998	Cross-sectional	7 healthy and 8 gingivitis 21 CP		BOP, PD, CAL, PI	GCF subgingival plaque	ICTP, 40 subgingival taxa	Positive relation between ICTP and putative pathogens
Perinetti et al. (2008)	Intervention	16 CP	15, 60 days after SRP	PD, PI, CAL, BOP	GCF	Alkaline phosphatase	Decreased on day 15, increased on day 60 after SRP
Pradeep et al. (2009a, b)	Intervention	20 healthy, 20 gingivitis, 20 CP	CP patients also 6-8 weeks after SRP	PD, CAL, GI	GCF	MCP-1	Increased gradually with severity of periodontal disease, decreased after SRP, correlated with clinical periodontal data; may be a biomarker

Pradeep et al. (2009a, b)	Intervention	20 healthy, 20 gingivitis, 20 CP	CP patients also 6–8 weeks after SRP		GCF	IL-18, MCP-1	Higher levels in CP, +correlation between IL-18 and MCP-1; IL-18 may induce MCP-1, promote inflammation
Rescala et al. (2010)	Cross-sectional	20 GCP, 17 GAgP, 10 gingivitis			GCF	IL-1 β , IL-2, IL-4, IL-8, elastase, IFN- γ	IL-1 β elastase were higher in deep sites than shallow sites in periodontitis groups. No significant difference between CP and AgP in GCF levels of biomarkers
Rosin et al. (1999)	Cross-sectional	140 subjects		PI, SBI, GCF flow	Saliva, GCF	Peroxidase, lysozyme	No significant correlation between clinical periodontal indices and salivary data
Srinath et al. (2010)	Cross-sectional	15 healthy, 15 gingivitis, 15 CP		GI, Russell periodontal index	GCF, saliva	Melatonin	Lowest in CP, may have a protective role in periodontal disease
Suma et al. (2009)	Cross-sectional	Healthy Gingivitis CP		PD	GCF, saliva	Lysozyme activity	Lysozyme activity correlated with PD
Teles et al. (2010a)	Cross-sectional	20 healthy, 20 CP			GCF; subgingival plaque	IL-1 β , IL-8, MMP-8, 40 bacterial taxa	+ correlation between clinical indices, GCF cytokines, orange and red complexes Healthy sites of CP had higher cytokine levels than healthy group
Thorat Manojkumar et al. (2010)	Intervention	20 healthy, 20 gingivitis, 20 CP	CP patients also 8 weeks after SRP	PD, CAL	GCF	Oncostatin M	Highest in CP, +correlation with PD, CAL; can be a biomarker in periodontitis
Toker et al. (2008)	Intervention	15 healthy, 15 generalized AgP	6 weeks after SRP	PD, CAL, PI, BOP, GI	GCF	IL-1 β , IL-1ra, IL-10	IL-1 β is high and decreased after SRP
Tsalikis et al. (2001)	Intervention	12 advanced periodontitis	4 weeks after SRP	PD, CAL, BOP	GCF	Aspartate amino transferase	AST decreased after SRP

(continued)

Table 12.9 (continued)

Reference	Study design	Study groups	Follow-up period	Clinical parameter	Biological sample	Biological parameter	Results
Türkoğlu et al. (2009)	Cross-sectional			PD, CAL, PI, BOP, PBI	GCF	IL-18, cathelicidin LL-37	Cathelicidin LL-37 increased in CP, may play a role in destruction
Tüter et al. (2007)	Cross-sectional	20 CP, 17 healthy			Serum, GCF	hsCRP	No difference between groups
Yetkin Ay et al. (2009)	Cross-sectional	40 CP, 20 healthy		GI, PD, CAL, BOP, PI	GCF	IL-11, IL-17	Healthy group had highest IL-11; IL-17 ratio
Yin et al. (2000)	Intervention	Periodontitis, gingivitis, healthy 7 CP	14 days after SRP		GCF	t-PA, PAI-2	Decreased after SRP, may be diagnostic markers
Yoshinari et al. (2004)	Intervention		Before and after SRP	Routine clinical parameters	GCF, gingiva	IL-1 α , IL-1 β , IL-1ra	IL-1 is effective for evaluating the gingival inflammation
Zheng et al. (2006)	Cross-sectional	21 CP, 19 gingivitis, 20 healthy controls			GCF, serum	PAF	PAF levels were higher in CP both in GCF and serum. Serum and GCF levels correlated also with clinical parameters. PAF may play a role in pathogenesis

CP chronic periodontitis, GCF gingival crevicular fluid, SOD superoxide dismutase, LPO lipid peroxidation, TOS total oxidant status, t-PA the tissue/blood vessel type plasminogen activator, u-PA urokinase type plasminogen activator, PAI plasminogen activator inhibitor, PA plasminogen activators, TAOC total antioxidant capacity, SRP scaling and root planing, PAF platelet activating factor, MCP-3 monocyte chemoattractant protein-3, IFN- γ interferon- γ , GAgP generalized aggressive periodontitis, PD probing depth, CAL clinical attachment level, BOP bleeding on probing, PI plaque index, IL interleukin, CRP C-reactive protein, SDD subantimicrobial dose doxycycline, ICTP carboxy-terminal telopeptide of type I collagen, GI gingival index, MMP matrix metalloproteinase, GSH reduced glutathione, GSSG oxidized glutathione, PAR₂ protease activated receptor 2, TNF tumour necrosis factor, DM diabetes mellitus, PGE₂ prostaglandin E₂, RANKL receptor activator of nuclear factor κ B ligand, MCP-1 monocyte chemoattractant protein-1, TACE transcatheter arterial chemoembolization, SBI sulcus bleeding index, IL-1ra interleukin-1 receptor antagonist, AST aspartate amino transferase, PBI papilla bleeding index, hsCRP high-sensitive C-reactive protein

for components of the gingival crevice fluid with a potential diagnostic value for periodontitis. They have evaluated almost 100 components of gingival crevice fluid with regard to their potential utility to detect a case of periodontitis, that is, to distinguish periodontitis from health and gingivitis; to classify a case of periodontitis, that is, chronic periodontitis or aggressive periodontitis; to plan treatment for the patient on the basis of the level of disease activity; and to monitor the treated patient based on the level of disease activity (Loos and Tjoa 2005). The reader may also refer to several other useful reviews on the topic of gingival crevice fluid and related markers (Lamster 1997; Champagne et al. 2003; Delima and Van Dyke 2003; Ebersole 2003; Eley and Cox 2003; Giannobile et al. 2003; Goodson 2003; Pollanen et al. 2003; Uitto et al. 2003; Armitage 2003; Ozmeric 2004; Heitz-Mayfield 2005; Taba et al. 2005; Kinney et al. 2007; Lamster and Ahlo 2007; Gorr 2009; Buduneli and Kinane 2011; Gorr and Abdolhosseini 2011).

12.5.3.5 Commercial Diagnostic Kits Based on GCF Proteolytic and Hydrolytic Enzyme Levels

MMPs are a family of structurally related but genetically distinct enzymes that degrade extracellular matrix (ECM) and basement membrane (BM) components. This group of 23 human enzymes is classified into collagenases, gelatinases, stromelysins, membrane-type MMPs, and other MMPs, mainly based on the substrate specificity and molecular structure. MMPs are involved in physiological processes such as tissue development, remodelling, and wound healing and play important roles in the regulation of cellular communication, molecular shedding, and immune functions by processing bioactive molecules including cell surface receptors, cytokines, hormones, defensins, adhesion molecules, and growth factors. MMP activity is controlled by changes in the delicate balance between the expression and synthesis of MMPs and their major endogenous inhibitors, tissue inhibitors of matrix metalloproteinases (TIMPs). The catalytic competence of MMPs is controlled through the activation of proenzymes and the inhibition of

the activation or activity by TIMPs (Reynolds and Meikle 1997; Uitto et al. 2003; Eley and Cox 2003; Sorsa et al. 2004, 2006; Buduneli and Kinane 2011).

Collagenase-2 (MMP-8), also called neutrophil collagenase, has previously been thought to be expressed solely by leukocytes of the polymorphonuclear (PMN) lineage cell line. MMP-8 is synthesized during the maturation of PMNs in the bone marrow; it then becomes glycosylated and is pre-stored in the subcellular-specific granules, from where it can be released by selective degranulation in response to appropriate triggering stimuli. The neutrophil (PMN)-type MMP-8 is more glycosylated than other MMPs including the fibroblast-type MMP-8 (11,15); these carbohydrate moieties of PMN-type MMP-8 are believed to act as targeting signals directing its subcellular location into intracellular-specific granules of PMNs and may explain the sensitivity of PMN-type MMP-8 to activation and inactivation by reactive oxygen species. MMP-8 can also be produced (de novo expressed) by articular chondrocytes and during various inflammatory diseases such as bronchitis, asthma, periodontitis and arthritis by resident synovial and gingival fibroblasts, epithelial cells/keratinocytes, odontoblasts, oral cancer cells, monocyte/macrophages, and plasma cells. Like MMP-1, MMP-8 also has wide substrate specificity, and its activities overlap with the other collagenases (MMP-1 and MMP-13). Recently, MMP-8 has been shown to exert unexpected anti-inflammatory defensive and protective characteristics against the spread of experimental skin cancer and inflammatory lung diseases probably by processing anti-inflammatory cytokines and chemokines as well as regulating inflammatory cell apoptosis and immune response (Sorsa et al. 2006 and references therein; Sorsa et al. 2011).

Several studies strongly indicate that the amount and activity of MMP-8 are highly increased in gingival crevicular fluid (GCF) of diseased periodontal pockets in chronic periodontitis patients. The increased amount and activity of MMP-8 correlate with the severity of periodontal disease (Pozo et al. 2005; Passoja et al. 2008; Xu et al. 2008) and peri-implantitis (Kivelä-Rajamäki et al. 2003a; Kivelä-Rajamäki et al.

2003b; Xu et al. 2008). GCF levels of MMP-8 are also related to progression episodes and treatment responses in patients with chronic periodontitis (Hernández et al. 2010). Recently, Reinhardt et al. (2010) showed that elevated GCF biomarkers of inflammation and bone resorption from a small number of moderate/deep sites have the potential to identify patients who are vulnerable to progressive periodontitis, and SDD may modify that risk. The levels of MMP-8 were significantly higher in GCF when *Prevotella intermedia*, *Tannerella forsythia*, and *Treponema denticola* were present (Söder et al. 2006; Jacob et al. 2012). Type 2 diabetes mellitus does not seem to significantly affect GCF levels of MMP-8 (Kardeşler et al. 2010). When gingival crevicular fluid (GCF) levels of matrix metalloproteinase MMP-8 and MMP-13 and tissue inhibitor of MMP (TIMP)-1 in patients with rheumatoid arthritis (RA) and systemically healthy counterparts with inflammatory periodontal disease were compared, it was shown that the total amounts of MMP-8 were lower in the healthy control group than in RA–gingivitis, RA–periodontitis, and healthy–periodontitis groups ($P < 0.05$). Patients with RA and gingivitis or periodontitis exhibited levels of MMP-8 that were similar to systemically healthy counterparts (Biyikoğlu et al. 2009).

Periodontal treatment such as scaling and root planing has been documented to decrease MMP-8 levels and activity markedly (Kinane et al. 2003). Periodontal pockets or sites with poor response to treatment have been documented to have persistently elevated MMP-8 levels (Mäntylä et al. 2006; Hernández et al. 2010). Several human and animal studies have demonstrated that the elevated MMP-8 associated with periodontal and inflammatory tissue destruction is sensitive to inhibition and reduction by subantimicrobial dose doxycycline medication (SDD) (Golub et al. 1990, 1997, 2008; Emingil et al. 2004; Lee et al. 2004; Ağan et al. 2006; Reinhardt et al. 2010).

Recently, monoclonal antibodies for MMP-8 were developed to be utilized in a **chair-side dipstick test for MMP-8** that allowed the development of a novel sensitive, specific, rapid, and practical immunological chairside dipstick test for MMP-8 in GCF and peri-implant sulcular fluid (PISF) (Sorsa et al. 1999, 2004). The test, bearing

resemblance to pregnancy home test kits (Fig. 12.54), can be performed by a dentist without specific equipment and measures the GCF MMP-8 level in 5 min (Fig. 12.55) (Mäntylä et al. 2003).

The test stick results were found to be in good agreement with the quantitative MMP-8 immunofluorometric assay (IFMA) analysis from GCF collected from periodontally healthy individuals and from gingivitis and periodontitis patients when all positive test results (+ to +++) were analysed. The agreement was very good when the Kappa-statistics ($k = 0.81$) were calculated, and the chairside GCF MMP-8 test provided a sensitivity of 0.83 and specificity of 0.96 (Mäntylä et al. 2003).

It differentiates periodontitis from gingivitis and healthy control sites, and reduction of GCF MMP-8 levels can be observed after successful periodontal treatment (Mäntylä et al. 2003). The mean MMP-8 concentrations, as determined with the chairside test, in smokers were lower than in nonsmokers, but in smokers' and nonsmokers' sites with progressive disease, MMP-8 concentrations were similar. Sites with exceptionally elevated MMP-8 concentrations were clustered in smokers who also showed a poor response to SRP. In these sites, the MMP-8 concentration did not decrease with SRP, and these sites were easily identified by the MMP-8 test (Mäntylä et al. 2006). GCF MMP-8 level testing is a very useful tool to monitor the beneficial effects of adjunctive subantimicrobial doxycycline medication for periodontitis patients (Emingil et al. 2004). This rapid point-of-care test developed for periodontitis is obviously a useful tool also for monitoring of peri-implantitis (Kivelä-Rajamäki et al. 2003a, b; Xu et al. 2008).

The **DentoAnalyzer** is designed as a portable user-friendly benchtop instrument. It automatically conducts the whole assay process, that is, steps like liquid handling as well as readout based on a software programme and a robust algorithm. The key component is a cartridge consisting (1) of a liquid-handling module containing all relevant reagents for immunological reactions like clinical sample, conjugate, wash buffers, and substrate and (2) a reaction chamber containing six filters including positive and negative controls, where the immunological reactions take place (Munjal et al. 2007a).

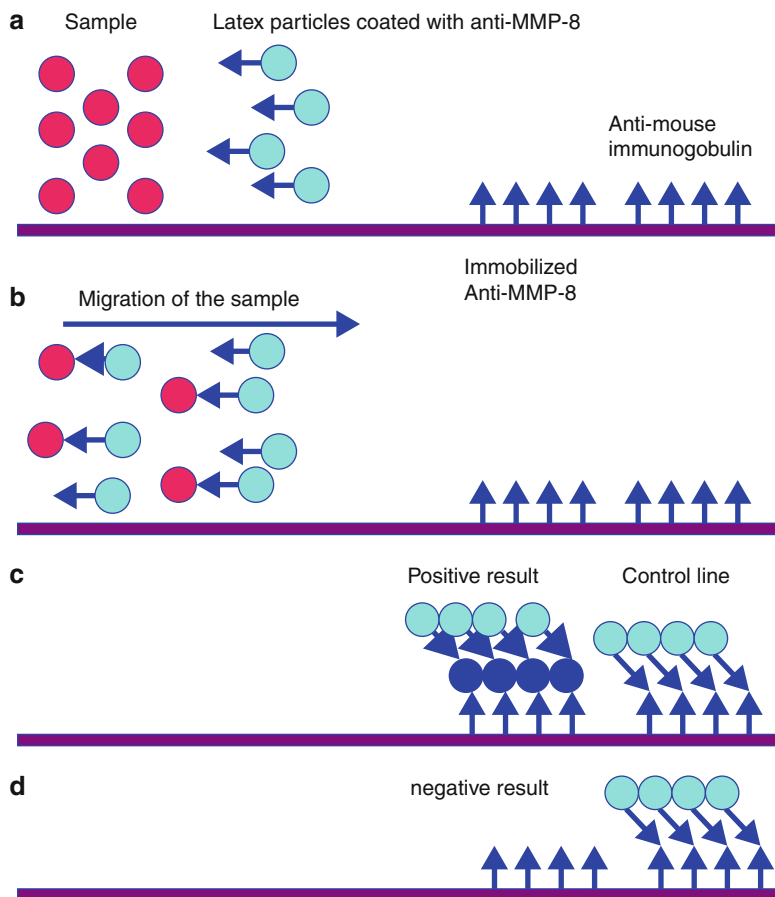


Fig. 12.54 Principle of the immunochromathographic MMP-8 dip-stick chair-side test in gingival crevicular fluid (GCF), peri-implant sulcular fluid (PISF) and other oral fluids (i.e. mouth-rinse samples, saliva, etc.). The MMP-8 test stick is based on the immunochromatography principle that uses two monoclonal antibodies specific for different epitopes of MMP-8. One is immobilized onto a nitrocellulose membrane to form a catching zone and the other onto blue latex particles. When the sample collected from the gingival crevice with a strip is diluted into the test kit buffer, the tip of an MMP-8 test stick is then immersed in the buffer for about 10 s. Fluid is absorbed into the stick, and MMP-8 in buffer migrates along the test stick and binds to the blue latex-labelled MMP-8 antibody on the stick (*phase a*). The complexes then migrate with the sample fluid across the nitrocellulose membrane

over the catching zone containing the other MMP-8 antibody (*phase b*). A sufficient MMP-8 concentration in the sample results in a blue line within 5 min in this zone when MMP-8 carrying latex is bound to it (*phase c*). The stick also contains latex-labelled mouse immunoglobulin, that after migration attaches to the anti-mouse immunoglobulin zone, demonstrating two blue lines to indicate successful performance of the test (*phase c*). When the MMP-8 concentration in the sample is below the detection level, only one (control) blue line appears, indicating a successfully performed test with negative result (*phase d*). The immunoassay can be adapted for any MMPs, TIMPs, PMN elastase, PMN lipocalin etc. alone and/or in any combinations with tissue degradation products (Sorsa et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

The test results from the analyzer were compared with quantitative MMP-8 immunofluorometric assay (IFMA) and in-house enzyme-linked immunosorbent assay (ELISA) as well as with the periodontal state. Preliminary results of analyzer measurements of clinical samples

showed a good agreement with the results from IFMA and in-house ELISA and with the clinical picture (Munjali et al. 2007b). Sorsa et al. (2010) compared four different methods for GCF sample MMP-8 detection in site-specific periodontal diagnostics: a time-resolved immunofluorometric

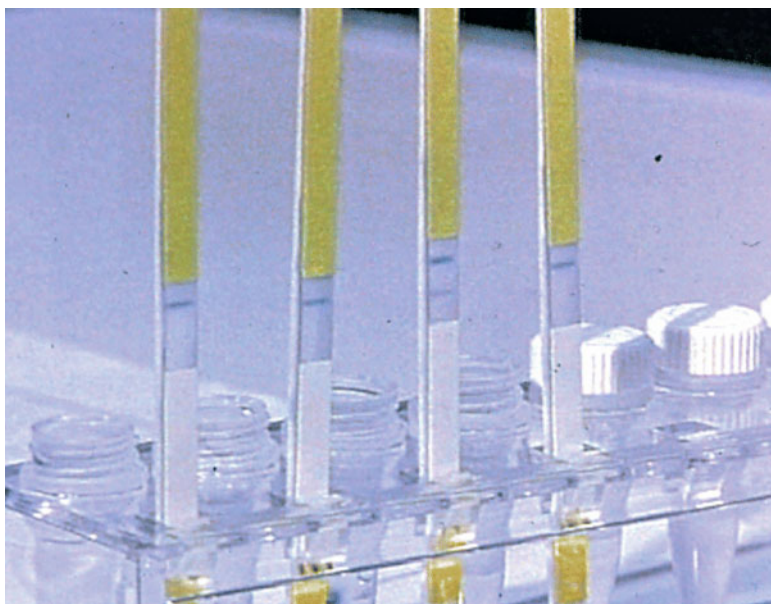


Fig. 12.55 Demonstration of the matrix metalloproteinase-8 (MMP-8)-specific test stick for the chair-side use. Prior to testing, the crown of the tooth is gently cleaned, dried and isolated with cotton rolls to exclude contamination with saliva. The tip of a filter-paper sampling strip for absorbing GCF is gently placed into the gingival sulcus for 30 s, avoiding any bleeding from the marginal gingiva. The sampling strip is then placed into a test tube containing 0.5 ml of HEPES-buffer, pH 7.4. The proteins absorbed by the sampling strip are then eluted into the buffer for 5 min. The tip of an MMP-8 test stick is then immersed in the buffer for about 10 s. Fluid is absorbed into the stick and MMP-8 in the buffer migrates along the test stick and binds to the blue antibody labeled latex particles on the

stick. The blue latex particles then migrate with the sample fluid across the nitrocellulose membrane over the catching zone containing the other MMP-8 antibody. A sufficient MMP-8 concentration in the sample results in a blue line in this zone when MMP-8 carrying latex is bound to it. The test is positive if the *blue line* appears in 5 min. The test was designed to give detection limit of 1 mg/l based on preliminary studies on MMP-8 in GCF from periodontitis, gingivitis and healthy sites. The appearance of a second blue line in the detection area indicates a gingival crevicular fluid (GCF) MMP-8 concentration of $\geq 1,000 \mu\text{g/l}$ (Mäntylä et al. 2003, 2006. Reprinted with permission from John Wiley & Sons, Inc.)

assay (IFMA), an MMP-8-specific chairside dipstick test, a device for rapid MMP-8 detection applying a sandwich-based immunoassay system using ABICAP filters (DentoAnalyzer by Dentognostics GmbH, Jena, Germany), and periodontitis patients' samples were evaluated also by a commercially available ELISA kit by Amersham. The results showed that immunofluorometric assay and DentoAnalyzer can detect MMP-8 from GCF samples and these methods are comparable. Using Western immunoblot, it was confirmed that IFMA and DentoAnalyzer can detect activated 55-kDa MMP-8 species especially in periodontitis-affected GCF.

It has been showed that the assessment of an MMP-8 in GCF with this noninvasive method is

able to assess and monitor the pathophysiological status of the periodontium (Sorsa et al. 2010). The determination of an MMP-8 with the *DentoAnalyzer* enables the dentist to distinguish between the healthy and the periodontitis-affected sites within 20 min (Prescher et al. 2007).

References

- Aas JA, Barbuto SM, Alpagot T, Olsen I, Dewhirst FE, Paster BJ. Subgingival plaque microbiota in HIV+ patients. *J Clin Periodontol*. 2007;34:189–95.
- Abbas F, Hart AA, Oosting J, van der Velden U. Effect of training and probing force on the reproducibility of pocket depth measurements. *J Periodontal Res*. 1982;17:226–34.
- Abiko Y, Sato T, Mayanagi G, Takahashi N. Profiling of subgingival plaque biofilm microflora from periodon-

- tally healthy subjects and from subjects with periodontitis using quantitative real-time PCR. *J Periodontol Res.* 2010;45:389–95.
- Aemaimanan P, Sattayasai N, Wara-aswapati N, Pitiphat W, Suwannarong W, Prajaneh S, Taweekaisupapong S. Alanine aminopeptidase and dipeptidyl peptidase IV in saliva of chronic periodontitis patients. *J Periodontol.* 2009;80:1809–14.
- Ağan S, Sönmez S, Serdar M. The effect of topical doxycycline usage on gingival crevicular fluid MMP-8 levels of chronic and aggressive periodontitis patients: a pilot study. *Int J Dent Hyg.* 2006;4:114–21.
- Agrawal P, Sanikop S, Patil S. New developments in tools for periodontal diagnosis. *Int Dent J.* 2012;62:57–64.
- Ai J, Smith B, Wong DT. Saliva Ontology: an ontology-based framework for a Salivaomics Knowledge Base. *BMC Bioinformatics.* 2010;11:302.
- Akalin FA, Toklu E, Renda N. Analysis of superoxide dismutase activity levels in gingiva and gingival crevicular fluid in patients with chronic periodontitis and periodontally healthy controls. *J Clin Periodontol.* 2005;32:238–43.
- Akesson L, Håkansson J, Rohlin M. Comparison of panoramic and intraoral radiography and pocket probing for the measurement of the marginal bone level. *J Clin Periodontol.* 1992;19:326–32.
- Akalin FA, Baltacıoğlu E, Alver A, Karabulut E. Lipid peroxidation levels and total oxidant status in serum, saliva and gingival crevicular fluid in patients with chronic periodontitis. *J Clin Periodontol.* 2007;34:558–65.
- Al-Ahmad A, Wunder A, Auschill TM, Follo M, Braun G, Hellwig E, Arweiler NB. The in vivo dynamics of *Streptococcus* spp, *Actinomyces naeslundii*, *Fusobacterium nucleatum* and *Veillonella* spp. in dental plaque biofilm as analysed by five colour multiplex fluorescence in situ hybridization. *J Med Microbiol.* 2007;56:681–7.
- Al-Ahmad A, Follo M, Selzer AC, Hellwig E, Hannig M, Hannig C. Bacterial colonization of enamel in situ investigated using fluorescence in situ hybridization. *J Med Microbiol.* 2009;58:1359–66.
- Al-Ahmad A, Roth D, Wolkewitz M, Wiedmann-Al-Ahmad M, Follo M, Ratka-Krüger P, Deimling D, Hellwig E, Hannig C. Change in diet and oral hygiene over an eight week period: effects on oral health and oral biofilm. *Clin Oral Investig.* 2010;14:391–6.
- Albandar JM, Abbas DK. Radiographic quantification of alveolar bone level changes. Comparison of 3 currently used methods. *J Clin Periodontol.* 1986;13:810–3.
- Al-Ghamdi HS, Anil S. Serum antibody levels in smoker and non-smoker Saudi subjects with chronic periodontitis. *J Periodontol.* 2007;78:1043–50.
- Ali M, Parlapalli R. Signal processing overview of optical coherence tomography systems for medical imaging. *SPRAB9.* 2010;1–22.
- Allaker RP, Aduse-Opoku J, Batten JE, Curtis MA. Natural variation within the principal arginine-specific protease gene, *prpR1*, of *Porphyromonas gingivalis*. *Oral Microbiol Immunol.* 1997;12:298–302.
- Alpha CX, Guthmiller JM, Cummings HE, Schomberg LL, Noorani SM. Molecular analysis of *Peptostreptococcus micros* isolates from patients with periodontitis. *J Periodontol.* 2001;72:877–82.
- Alves Rde V, Machion L, Andia DC, Casati MZ, Sallum AW, Sallum EA. Reproducibility of clinical attachment level and probing depth of a manual probe and a computerized electronic probe. *J Int Acad Periodontol.* 2005;7:27–30.
- Amann RI. In situ identification of microorganisms by whole cell hybridization with rRNA-targeted nucleic acid probes. In: Akkermans ADL, van Elsas JD, de Bruijn FJ, editors. *Molecular microbial ecology manual*, 3.3.6. Dordrecht: Kluwer; 1995. p. 1–15.
- Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol.* 1990;56:1919–25.
- Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev.* 1995;59:143–69.
- Amarasena N, Yoshihara A, Hirotomi T, Takano N, Miyazaki H. Association between serum calcium and periodontal disease progression in non-institutionalized elderly. *Gerodontology.* 2008;25:245–50.
- American Academy of Periodontology. Proceedings of the world workshop in clinical periodontics. Chicago: American Academy of Periodontology; 1989. p. 1–13.
- Armitage GC. Periodontal diseases: diagnosis. *Ann Periodontol.* 1996;1:37–215.
- Armitage GC, Research, Science and Therapy Committee of the American Academy of Periodontology. Diagnosis of periodontal diseases. *J Periodontol.* 2003;74:1237–47.
- Armitage GC, Svanberg GK, Löe H. Microscopic evaluation of clinical measurements of connective tissue attachment levels. *J Clin Periodontol.* 1977;4:173–90.
- Armitage GC, Jeffcoat MK, Chadwick DE, Taggart Jr EJ, Numabe Y, Landis JR, Weaver SL, Sharp TJ. Longitudinal evaluation of elastase as a marker for the progression of periodontitis. *J Periodontol.* 1994;65:120–8.
- Asai Y, Jinno T, Igarashi H, Ohyama Y, Ogawa T. Detection and quantification of oral treponemes in subgingival plaque by real-time PCR. *J Clin Microbiol.* 2002;40:3334–40.
- Asikainen S, Chen C. Oral ecology and person-to-person transmission of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. *Periodontol* 2000. 1999;20:65–81.
- Asikainen S, Alaluusua S, Saxen L. Recovery of actinomycescomitans from teeth, tongue, and saliva. *J Periodontol.* 1991;62:203–6.
- Asikainen S, Chen C, Slots J. *Actinobacillus actinomycetemcomitans* genotypes in relation to serotypes and periodontal status. *Oral Microbiol Immunol.* 1995;10:65–8.
- Atassi F, Newman HN, Bulman JS. Probe tine diameter and probing depth. *J Clin Periodontol.* 1992;19:301–4.

- Atieh MA. Accuracy of real-time polymerase chain reaction versus anaerobic culture in detection of *Aggregatibacter actinomycetemcomitans* and *Porphyromonas gingivalis*: a meta-analysis. J Periodontol. 2008;79:1620–9.
- Atkinson JC, Dawes C, Ericson T, Fox PC, Gandara BK, Malamud D, Mandel ID, Navazesh M, Tabak LA. Guidelines for saliva nomenclature and collection. Ann N Y Acad Sci. 1993;694:Guidelines for saliva nomenclature and collection—xii.
- Aurer A, Aurer-Kozelj J, Stavljenić-Rukavina A, Kalenić S, Ivić-Kardum M, Haban V. Inflammatory mediators in saliva of patients with rapidly progressive periodontitis during war stress induced incidence increase. Coll Antropol. 1999;23:117–24.
- Badersten A, Nilveus R, Egelberg J. Effect of nonsurgical periodontal therapy. VII. Bleeding, suppuration and probing depth in sites with probing attachment loss. J Clin Periodontol. 1985;12:432–40.
- Baek JH, Na J, Lee BH, Choi E, Son WS. Optical approach to the periodontal ligament under orthodontic tooth movement: a preliminary study with optical coherence tomography. Am J Orthod Dentofacial Orthop. 2009;135:252–9.
- Baker PJ, Butler R, Wikesjö UM. Bacterial sampling by absorbent paper points, an in vitro study. J Periodontol. 1991;62:142–6.
- Baltacıoğlu E, Akalin FA, Alver A, Değer O, Karabulut E. Protein carbonyl levels in serum and gingival crevicular fluid in patients with chronic periodontitis. Arch Oral Biol. 2008;53:716–22.
- Barbiera F, Bartolotta TV, Lo Casto A, Pardo S, Rossello M, De Maria M. Intra-articular osteoid osteoma: diagnostic imaging in three cases. Radiol Med (Torino). 2002;103:464–73.
- Barendregt DS, Van der Velden U, Reiker J, Loos BG. Clinical evaluation of tine shape of 3 periodontal probes using 2 probing forces. J Clin Periodontol. 1996;23:397–402.
- Barnfather KD, Cope GF, Chapple IL. Effect of incorporating a 10 minute point of care test for salivary nicotine metabolites into a general practice based smoking cessation programme: randomised controlled trial. BMJ. 2005;331:999.
- Baron AC, Gansky SA, Ryder MI, Featherstone JD. Cysteine protease inhibitory activity and levels of salivary cystatins in whole saliva of periodontally diseased patients. J Periodontol Res. 1999;34:437–44.
- Barton DJ, Clark SJ, Eleazer PD, Scheetz JP, Farman AG. Tuned-aperture computed tomography versus parallax analog and digital radiographic images in detecting second mesiobuccal canals in maxillary first molars. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2003;96:223–8.
- Baum BJ, Yates 3rd JR, Srivastava S, Wong DT, Melvin JE. Scientific frontiers: emerging technologies for salivary diagnostics. Adv Dent Res. 2011;23:360–8.
- Behle JH, Sedaghatfar MH, Demmer RT, Wolf DL, Celenti R, Klebschull M, Belusko PB, Herrera-Abreu M, Lalla E, Papapanou PN. Heterogeneity of systemic inflammatory responses to periodontal therapy. J Clinical Periodontol. 2009;36:287–94.
- Beikler T, Peters U, Ehmke B, Flemmig TF. Sequence analysis of *kgp* in *Porphyromonas gingivalis* isolates from periodontitis patients. Oral Microbiol Immunol. 2003a;18:393–7.
- Beikler T, Peters U, Prajane S, Prior K, Ehmke B, Flemmig TF. Prevalence of *Porphyromonas gingivalis* *fimA* genotypes in Caucasians. Eur J Oral Sci. 2003b;111:390–4.
- Beikler T, Karch H, Flemmig TF. Microbiologic diagnosis in periodontal therapy. Joint statement of the German Society of Periodontology (DGP) and the German Society of Dental, Oral, and Maxillofacial Medicine (DGZMK) (in German). Dtsch Zahnärztl Z. 2005;60:660–2.
- Beikler T, Schnitzer S, Abdeen G, Ehmke B, Eisenacher M, Flemmig TF. Sampling strategy for intraoral detection of periodontal pathogens before and following periodontal therapy. J Periodontol. 2006;77:1323–32.
- Beuf O, Briguet A, Lissac M. In vitro magnetic resonance imaging of rodent teeth. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1997;84:582–5.
- Birek P, McCulloch CA, Hardy V. Gingival attachment level measurements with an automated periodontal probe. J Clin Periodontol. 1987;14:472–7.
- Biyyikoğlu B, Buduneli N, Kardeşler L, Aksu K, Pitkala M, Sorsa T. Gingival crevicular fluid MMP-8 and -13 and TIMP-1 levels in patients with rheumatoid arthritis and inflammatory periodontal disease. J Periodontol. 2009;80:1307–14.
- Björn H, Halling A, Thyberg H. Radiographic assessment of marginal bone loss. Odontol Revy. 1969;20:165–79.
- Bonta Y, Zambon JJ, Genco RJ, Neiders ME. Rapid identification of periodontal pathogens in subgingival plaque: comparison of indirect immunofluorescence microscopy with bacterial culture for detection of *Actinobacillus actinomycetemcomitans*. J Dent Res. 1985;64:793–8.
- Bottari B, Ercolini D, Gatti M, Neviani E. Application of FISH technology for microbiological analysis: current state and prospects. Appl Microbiol Biotechnol. 2006;73:485–94.
- Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CM, Savelkoul PH. Comparison of real-time PCR and culture for detection of *Porphyromonas gingivalis* in subgingival plaque samples. J Clin Microbiol. 2003;41:4950–4.
- Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CM, Savelkoul PH. Periodontal pathogens: a quantitative comparison of anaerobic culture and real-time PCR. FEMS Immunol Med Microbiol. 2005;45:191–9.
- Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CM, Savelkoul PH. The additional value of real-time PCR in the quantitative detection of periodontal pathogens. J Clin Periodontol. 2006;33:427–33.
- Boutaga K, Savelkoul PH, Winkel EG, van Winkelhoff AJ. Comparison of subgingival bacterial sampling with oral lavage for detection and quantification of periodontal pathogens by real-time polymerase chain reaction. J Periodontol. 2007;78:79–86.
- Boyer BP, Ryerson CC, Reynolds HS, Zambon JJ, Genco RJ, Snyder B. Colonization by *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis* and *Prevotella intermedia* in adult periodontitis patients as

- detected by the antibody-based evalusite test. *J Clin Periodontol.* 1996;23:477–84.
- Braga RR, Carvalho MA, Bruña-Romero O, Teixeira RE, Costa JE, Mendes EN, Farias LM, Magalhães PP. Quantification of five putative periodontal pathogens in female patients with and without chronic periodontitis by real-time polymerase chain reaction. *Anaerobe.* 2010;16:234–9.
- Brägger U. Radiographic parameters: biological significance and clinical use. *Periodontol* 2000. 2005;39:73–90.
- Brezinsky M. Optical coherence tomography. London: Elsevier; 2006.
- Buduneli N, Kinane DF. Host-derived diagnostic markers related to soft tissue destruction and bone degradation in periodontitis. *J Clin Periodontol.* 2011;38:85–105.
- Buduneli N, Buduneli E, Kardeşler L, Lappin D, Kinane DF. Plasminogen activator system in smokers and non-smokers with and without periodontal disease. *J Clin Periodontol.* 2005;32:417–24.
- Bustin SA. Developments in real-time PCR research and molecular diagnostics. *Expert Rev Mol Diagn.* 2010a;10:713–5.
- Bustin SA. Developments in real-time PCR research and molecular diagnostics. *Expert Rev Mol Diagn.* 2010;10:713–5.
- Bustin SA. Why the need for qPCR publication guidelines? The case for MIQE. *Methods.* 2010b;50:217–26.
- Casarin RC, Ribeiro Edel P, Mariano FS, Nociti Jr FH, Casati MZ, Gonçalves RB. Levels of *Aggregatibacter actinomycetemcomitans*, *Porphyromonas gingivalis*, inflammatory cytokines and species-specific immunoglobulin G in generalized aggressive and chronic periodontitis. *J Periodontol Res.* 2010;45:635–42.
- Casas A, Herrera A, D, Martín-Carnes J, González I, O'Connor A, Sanz M. Influence of sampling strategy on microbiologic results before and after periodontal treatment. *J Periodontol.* 2007;78:1103–12.
- Caton J, Greenstein G, Polson AM. Depth of periodontal probe penetration related to clinical and histologic signs of gingival inflammation. *J Periodontol.* 1981;52:626–9.
- Chairoungdua A, Smith DL, Pochard P, Hull M, Caplan MJ. Exosome release of beta-catenin: a novel mechanism that antagonizes Wnt signaling. *J Cell Biol.* 2010;190:1079–91.
- Chal-U-Dom O, Ludlow JB, Tyndall DA, Webber RL. Comparison of conventional and TACT>(Tuned Aperture Computed Tomography) digital subtraction radiography in detection of pericrestal bone-gain. *J Periodontol Res.* 2002;37:147–53.
- Champagne CM, Buchanan W, Reddy MS, Preisser JS, Beck JD, Offenbacher S. Potential for gingival crevice fluid measures as predictors of risk for periodontal diseases. *Periodontol* 2000. 2003;31:167–80.
- Chapple IL, Brock GR, Milward MR, Ling N, Matthews JB. Compromised GCF total antioxidant capacity in periodontitis: cause or effect? *J Clin Periodontol.* 2007;34:103–10.
- Chaves ES, Jeffcoat MK, Ryerson CC, Snyder B. Persistent bacterial colonization of *Porphyromonas gingivalis*, *Prevotella intermedia*, and *Actinobacillus actinomycetemcomitans* in periodontitis and its association with alveolar bone loss after 6 months of therapy. *J Clin Periodontol.* 2000;27:897–903.
- Chen C, Slots J. Microbiological tests for *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. *Periodontol* 2000. 1999;20:53–64.
- Chen H, Zheng P, Zhu H, Zhu J, Zhao L, El Mokhtari NE, Eberhard J, Lins M, Jepsen S. Platelet-activating factor levels of serum and gingival crevicular fluid in non-smoking patients with periodontitis and/or coronary heart disease. *Clin Oral Investig.* 2010;14:629–36.
- Chifor R, Hedeşiu M, Bolfa P, Catoi C, Crişan M, Serbănescu A, Badea AF, Moga I, Badea ME. The evaluation of 20 MHz ultrasonography, computed tomography scans as compared to direct microscopy for periodontal system assessment. *Med Ultrason.* 2011;13:120–6.
- Christersson LA, Fransson CL, Dunford RG, Zambon JJ. Subgingival distribution of periodontal pathogenic microorganisms in adult periodontitis. *J Periodontol.* 1992;63:418–25.
- Christodoulides N, Floriano PN, Miller CS, Ebersole JL, Mohanty S, Dharshan P, Griffin M, Lennart A, Ballard KL, King Jr CP, Langub MC, Kryscio RJ, Thomas MV, McDevitt JT. Lab-on-a-chip methods for point-of-care measurements of salivary biomarkers of periodontitis. *Ann N Y Acad Sci.* 2007;1098:411–28.
- Christgau M, Bader N, Felden A, Gradl J, Wenzel A, Schmalz G. Guided tissue regeneration in intrabony defects using an experimental bioresorbablepolydioxanone (PDS) membrane. A 24-month split-mouth study. *J Clin Periodontol.* 2002;29:710–23.
- Cohen L, Rams T, Slots J, Walker C. Independent analyses of microbiological samples by three testing laboratories. *J Dent Res.* 2001;80:219. Abstr. 1465.
- Colombo AV, da Silva CM, Haffajee A, Colombo AP. Identification of intracellular oral species within human crevicular epithelial cells from subjects with chronic periodontitis by fluorescence in situ hybridization. *J Periodontol Res.* 2007;42:236–43.
- Colombo APV, Boches SK, Cotton SL, Goodson JM, Kent R, Haffajee AD, Socransky SS, Hasturk H, van Dyke TE, Paster BJ. Comparisons of subgingival microbial profiles of refractory periodontitis, severe periodontitis and periodontal health using the Human Oral Microbe Identification Microarray (HOMIM). *J Periodontol.* 2009;80:1421–32.
- Combs DR, Reilly EA, Dawson 3rd DR, Avdiushko SA, Danaher RJ, Miller CS. Detection of human cytomegalovirus in dental plaque from individual periodontal sites by real-time polymerase chain reaction. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2008;106:840–4.
- Conrads G, Muttters R, Fischer J, et al. PCR reaction and dot-blot hybridization of oral pathogens within plaque samples of periodontally healthy individuals. *J Periodontol.* 1996;67:994–1003.
- Corbet EF, Ho DK, Lai SM. Radiographs in periodontal disease diagnosis and management. *Aust Dent J.* 2009;54 Suppl 1:S27–43.
- Correia FF, Plummer AR, Paster BJ, Dewhirst FE. Genome size of human oral *Treponema* species by

- pulsed-field gel electrophoresis. *Oral Microbiol Immunol.* 2004;19:129–31.
- Cotti E. Advanced techniques for detecting lesions in bone. *Dent Clin North Am.* 2010;54:215–35.
- Cotti E, Campisi G. Advanced radiographic techniques for the detection of lesions in bone. *Endod Top.* 2004;7:52–72.
- Cotti E, Campisi G, Garau V, Puddu G. A new technique for the study of periapical bone lesions: ultrasound real time imaging. *Int Endod J.* 2002;35:148–52.
- Cotti E, Campisi G, Ambu R, Dettori C. Ultrasound real-time imaging in the differential diagnosis of periapical lesions. *Int Endod J.* 2003;36:556–63.
- Cotti E, Simbola V, Dettori C, Campisi G. Echographic evaluation of bone lesions of endodontic origin: report of two cases in the same patient. *J Endod.* 2006;32:901–5.
- Cutando A, Galindo P, Gomez-Moreno G, Arana C, Bolanos J, Acuna-Castroviejo D, Wang HL. Relationship between salivary melatonin and severity of periodontal disease. *J Periodontol.* 2006;77:1533–8.
- D'Ercole S, Catamo G, Tripodi D, Piccolomini R. Comparison of culture methods and multiplex PCR for the detection of periodontopathogenic bacteria in biofilm associated with severe forms of periodontitis. *New Microbiol.* 2008;31:383–91.
- Dahlen G, Renvert S, Wikstrom M, Egelberg J. Reproducibility of microbiological samples from periodontal pockets. *J Clin Periodontol.* 1990;17:73–7.
- De La Garza-Ramos MA, Galán-Wong LJ, Caffesse RG, González-Salazar F, Pereyra-Alferez B. Detection of *Porphyromonas gingivalis* and *Streptococcus intermedius* in chronic periodontitis patients by multiplex PCR. *Acta Odontol Latinoam.* 2008;21:163–7.
- De la Pena VA, Diz Dios P, Tojo Sierra R. Relationship between lactate dehydrogenase activity in saliva and oral health status. *Arch Oral Biol.* 2007;52:911–5.
- de Queiroz AC, Taba Jr M, O'Connell PA, da Nóbrega PB, Costa PP, Kawata VK, Trevisan GL, Novaes Jr AB, de Souza SL, Palioto DB, Grisi MF. Inflammation markers in healthy and periodontitis patients: a preliminary data screening. *Braz Dent J.* 2008;19:3–8.
- Deinzer R, Mossanen BS, Herforth A. Methodological considerations in the assessment of gingival crevicular fluid volume. *J Clin Periodontol.* 2000;27:481–8.
- Delima AJ, Van Dyke TE. Origin and function of the cellular components in gingival crevice fluid. *Periodontol* 2000. 2003;31:55–76.
- Dezerega A, Pozo P, Hernández M, Oyarzún A, Rivera O, Dutzan N, Gutiérrez-Fernández A, Overall CM, Garrido M, Alcota M, Ortiz E, Gamonal J. Chemokine monocyte chemoattractant protein-3 in progressive periodontal lesions in patients with chronic periodontitis. *J Periodontol.* 2010;81:267–76.
- Diab-Ladki R, Pellat B, Chahine R. Decrease in the total antioxidant activity of saliva in patients with periodontal diseases. *Clin Oral Investig.* 2003;7:103–7.
- Dibart S, Chapple ILC, Skobe Z, Shusterman S, Needleman HL. Microbiological findings in prepubertal periodontitis. A case report. *J Periodontol.* 1998;69:1172–5.
- Dietrich T, Bernimoulin JP, Glynn RJ. The effect of cigarette smoking on gingival bleeding. *J Periodontol.* 2004;75:16–22.
- Dige I, Nilsson H, Kilian M, Nyvad B. In situ identification of streptococci and other bacteria in initial dental biofilm by confocal laser scanning microscopy and fluorescence in situ hybridization. *Eur J Oral Sci.* 2007;115:459–67.
- Dige I, Raarup MK, Nyengaard JR, Kilian M, Nyvad B. *Actinomyces naeslundii* in initial dental biofilm formation. *Microbiology.* 2009;155:2116–26.
- Dinsdale CR, Rawlinson A, Walsh TF. Subgingival temperature in smokers and non-smokers with periodontal disease. *J Clin Periodontol.* 1997;24:761–6.
- Dogan B, Saarela M, Asikainen S. Genotyping of *Actinobacillus actinomycetemcomitans* serotype d isolates based on polymerase chain reaction. *Oral Microbiol Immunol.* 1999a;14:387–90.
- Dogan B, Saarela MH, Jousimies-Somer H, Alaluusua S, Asikainen S. *Actinobacillus actinomycetemcomitans* serotype e – biotypes, genetic diversity and distribution in relation to periodontal status. *Oral Microbiol Immunol.* 1999b;14:98–103.
- Draper C. Advances in technology and periodontal therapy. *J Calif Dent Hyg Assoc.* 2010;25:12–4.
- Drexler W, Fujimoto JG, editors. Optical coherence tomography: technology and applications. New York: Springer; 2008.
- Duarte PM, da Rocha M, Sampaio E, Mestnik MJ, Feres M, Figueiredo LC, Bastos MF, Faveri M. Serum levels of cytokines in subjects with generalized chronic and aggressive periodontitis before and after non-surgical periodontal therapy: a pilot study. *J Periodontol.* 2010;81:1056–63.
- Dutzan N, Gamonal J, Silva A, Sanz M, Vernal R. Overexpression of forkhead box P3 and its association with receptor activator of nuclear factor-kappa B ligand, interleukin (IL) -17, IL-10 and transforming growth factor-beta during the progression of chronic periodontitis. *J Clin Periodontol.* 2009;36:396–403.
- Eberhard J, Menzel N, Dommisch H, Winter J, Jepsen S, Mutters R. The stage of native biofilm formation determines the gene expression of human beta-defensin-2, psoriasin, ribonuclease 7 and inflammatory mediators: a novel approach for stimulation of keratinocytes with in situ formed biofilms. *Oral Microbiol Immunol.* 2008;23:21–8.
- Ebersole JL. Humoral immune responses in gingival crevice fluid: local and systemic implications. *Periodontol* 2000. 2003;31:135–66.
- Eggesbo HB, Dolvik S, Stiris M, Sovik S, Storrosten OT, Kolmannskog F. Complementary role of MR imaging of ethmoidal sinus disease depicted at CT in cystic fibrosis. *Acta Radiol.* 2001;42:144–50.
- Ehmke B, Schmidt H, Beikler T, Kopp C, Karch H, Klaiber B, Flemmig TF. Clonal infection with *Actinobacillus actinomycetemcomitans* following periodontal therapy. *J Dent Res.* 1999;78:1518–24.
- Eick S, Pfister W. Comparison of microbial cultivation and a commercial nucleic acid based method for

- detection of periodontopathogenic species in subgingival plaque samples. *J Clin Periodontol.* 2002;29:638–44.
- Eley BM, Cox SW. Proteolytic and hydrolytic enzymes from putative periodontal pathogens: characterization, molecular genetics, effects on host defenses and tissues and detection in gingival crevice fluid. *Periodontol* 2000. 2003;31:105–24.
- Emingil G, Atilla G, Sorsa T, Luoto H, Kirilmaz L, Baylas H. The effect of adjunctive low-dose doxycycline therapy on clinical parameters and gingival crevicular fluid matrix metalloproteinase-8 levels in chronic periodontitis. *J Periodontol.* 2004;75:106–15.
- Emingil G, Kuula H, Sorsa T, Atilla G. Gingival crevicular fluid matrix metalloproteinase-25 and -26 levels in periodontal disease. *J Periodontol.* 2006;77:664–71.
- Emmerling H, Standley E. Probing into probes. Measuring the choices. *J Calif Dent Hyg Assoc.* 2010;25:12–4.
- Estrela CR, Pimenta FC, Alencar AH, Ruiz LF, Estrela C. Detection of selected bacterial species in intraoral sites of patients with chronic periodontitis using multiplex polymerase chain reaction. *J Appl Oral Sci.* 2010;18:426–31.
- Fedi Jr PF, Killoy WJ. Temperature differences at periodontal sites in health and disease. *J Periodontol.* 1992;63:24–7.
- Fitzsimmons TR, Sanders AE, Bartold PM, Slade GD. Local and systemic biomarkers in gingival crevicular fluid increase odds of periodontitis. *J Clin Periodontol.* 2010;37:30–6.
- Florianio PN, Christodoulides N, Miller CS, Ebersole JL, Spertus J, Rose BG, Kinane DF, Novak MJ, Steinhilb S, Acosta S, Mohanty S, Dharshan P, Yeh CK, Redding S, Furmaga W, McDevitt JT. Use of saliva-based nanobiochip tests for acute myocardial infarction at the point of care: a feasibility study. *Clin Chem.* 2009;55:1530–8.
- Ford PJ, Gamonal J, Seymour GJ. Immunological differences and similarities between chronic periodontitis and aggressive periodontitis. *Periodontol* 2000. 2010;53:111–23.
- Fox PC. Saliva composition and its importance in dental health. *Compend Contin Educ Dent.* 1989;10:S457–60.
- Furugen R, Hayashida H, Yamaguchi N, Yoshihara A, Ogawa H, Miyazaki H, Saito T. The relationship between periodontal condition and serum levels of resistin and adiponectin in elderly Japanese. *J Periodontal Res.* 2008a;43:556–62.
- Furugen R, Hayashida H, Yamaguchi N, Yoshihara A, Ogawa H, Miyazaki H, Saito T. The relationship between periodontal condition and serum levels of resistin and adiponectin in elderly Japanese. *J Periodontal Res.* 2008b;43:556–62.
- Gapski R, Hasturk H, Van Dyke TE, Oringer RJ, Wang S, Braun TM, Giannobile WV. Systemic MMP inhibition for periodontal wound repair: results of a multi-centre randomized-controlled clinical trial. *J Clin Periodontol.* 2009;36:149–56.
- García L, Tercero JC, Legido B, Ramos JA, Alemany J, Sanz M. Rapid detection of *Actinobacillus actinomycetemcomitans*, *Prevotella intermedia* and *Porphyromonas gingivalis* by multiplex PCR. *J Periodontal Res.* 1998;33:59–64.
- Garg G, Pradeep AR, Thorat MK. Effect of nonsurgical periodontal therapy on crevicular fluid levels of Cathepsin K in periodontitis. *Arch Oral Biol.* 2009;54:1046–51.
- Garito ML, Prihoda TJ, McManus LM. Salivary PAF levels correlate with the severity of periodontal inflammation. *J Dent Res.* 1995;74:1048–56.
- Garnick JJ, Silverstein L. Periodontal probing: probe tip diameter. *J Periodontol.* 2000;71:96–103.
- Garnick JJ, Keagle JG, Searle JR, King GE, Thompson WO. Gingival resistance to probing forces. II. The effect of inflammation and pressure on probe displacement in beagle dog gingivitis. *J Periodontol.* 1989;60:498–505.
- Gau V, Wong D. Oral fluid nanosensor test (OFNASET) with advanced electrochemical-based molecular analysis platform. *Ann N Y Acad Sci.* 2007;1098:401–10.
- Gaudino C, Cosgarea R, Heiland S, Csernus R, Beomonte Zobel B, Pham M, Kim TS, Bendszus M, Rohde S. MR-Imaging of teeth and periodontal apparatus: an experimental study comparing high-resolution MRI with MDCT and CBCT. *Eur Radiol.* 2011;21(12):2575–83. Epub 2011 Jul 31.
- Ge Z, Liu KZ, Xiang X, Yang Q, Hui J, Kohlenberg E, Sowa MG. Assessment of local hemodynamics in periodontal inflammation using optical spectroscopy. *J Periodontol.* 2011;82:1161–8.
- George KS, Reynolds MA, Falkler Jr WA. Arbitrarily primed polymerase chain reaction fingerprinting and clonal analysis of oral *Fusobacterium nucleatum* isolates. *Oral Microbiol Immunol.* 1997;12:219–26.
- Ghallab N, Shaker O. Salivary-soluble CD44 levels in smokers and non-smokers with chronic periodontitis: a pilot study. *J Periodontol.* 2010;81:710–7.
- Gheren LW, Cortell JR, Rodrigues E, Holzhausen M, Saad WA. Periodontal therapy reduces arginase activity in saliva of patients with chronic periodontitis. *Clin Oral Investig.* 2008;12:67–72.
- Giannobile WV. Periodontal surveillance – implications in the promotion of public health. *J Periodontol.* 2007;78:1177.
- Giannobile WV, Al-Shammari KF, Sarment DP. Matrix molecules and growth factors as indicators of periodontal disease activity. *Periodontol* 2000. 2003;31:125–34.
- Giannobile WV, Beikler T, Kinney JS, Ramseier CA, Morelli T, Wong DT. Saliva as a diagnostic tool for periodontal disease: current state and future directions. *Periodontol* 2000. 2009;50:52–64.
- Giannobile WV, McDevitt JT, Niedbala RS, Malamud D. Translational and clinical applications of salivary diagnostics. *Adv Dent Res.* 2011;23:375–80.
- Gibbs CH, Hirschfeld JW, Lee JG, Low SB, Magnusson I, Thousand RR, Yerneni P, Clark WB. Description and clinical evaluation of a new computerized periodontal probe – the Florida probe. *J Clin Periodontol.* 1988;15:137–44.

- Glas J, Beynon V, Bachstein B, Steckenbiller J, Manolis V, Euba A, Müller-Myhsok B, Folwaczny M. Increased plasma concentration of surfactant protein D in chronic periodontitis independent of SFTPD genotype: potential role as a biomarker. *Tissue Antigens*. 2008;72:21–8.
- Gmür R. Applicability of monoclonal antibodies to quantitatively monitor subgingival plaque for specific bacteria. *Oral Microbiol Immunol*. 1988;3:187–91.
- Gmür R, Thurnheer T. Direct quantitative differentiation between *Prevotella intermedia* and *Prevotella nigrescens* in clinical specimens. *Microbiology*. 2002;148(Pt 5):1379–87.
- Golub LM, Ciano S, Ramamamurthy NS, Leung M, McNamara TF. Low-dose doxycycline therapy: effect on gingival and crevicular fluid collagenase activity in humans. *J Periodontol Res*. 1990;25:321–30.
- Golub L, Lee HM, Greenwald RA, Ryan ME, Sorsa T, Salo T, et al. A matrix metalloproteinase inhibitor reduces bone-type collagen degradation fragments and specific collagenases in gingival crevicular fluid during adult periodontitis. *Inflamm Res*. 1997;46:310–9.
- Golub LM, Lee HM, Stoner JA, Sorsa T, Reinhardt RA, Wolff MS, et al. Subantimicrobial-dose doxycycline modulates gingival crevicular fluid biomarkers of periodontitis in postmenopausal osteopenic women. *J Periodontol*. 2008;79:1409–18.
- Golub LM, Lee HM, Stoner JA, Sorsa T, Reinhardt RA, Wolff MS, Ryan ME, Nummikoski PV, Payne JB. Subantimicrobial-dose doxycycline modulates gingival crevicular fluid biomarkers of periodontitis in postmenopausal osteopenic women. *J Periodontol*. 2008;79:1409–18.
- Gomes SC, Piccinin FB, Oppermann RV, Susin C, Marcantonio RA. The effect of smoking on gingival crevicular fluid volume during the treatment of gingivitis. *Acta Odontol Latinoam*. 2009;22:201–6.
- Gonçalves Lda R, Soares MR, Nogueira FC, Garcia C, Camisasca DR, Domont G, et al. Comparative proteomic analysis of whole saliva from chronic periodontitis patients. *J Proteomics*. 2010;73:1334–41.
- Goodson JM. Gingival crevice fluid flow. *Periodontol*. 2000. 2003;31:43–54.
- Goodson JM, Kondon N. Periodontal pocket depth measurements by fiber optic technology. *J Clin Dent*. 1988;1:35–8.
- Gorr SU. Antimicrobial peptides of the oral cavity. *Periodontol*. 2000. 2009;51:152–80.
- Gorr SU, Abdolhosseini M. Antimicrobial peptides and periodontal disease. *J Clin Periodontol*. 2011;38 Suppl 11:126–41.
- Grant MM, Brock GR, Matthews JB, Chapple IL. Crevicular fluid glutathione levels in periodontitis and the effect of non-surgical therapy. *J Clin Periodontol*. 2010;37:17–23.
- Greenstein G. The role of bleeding upon probing in the diagnosis of periodontal disease. A literature review. *J Periodontol*. 1984;55:684–8.
- Griffiths GS. Formation, collection and significance of gingival crevice fluid. *Periodontol*. 2000. 2003;31:32–42.
- Grisi MF, Novaes AB, Ito IY, Salvador SL. Relationship between clinical probing depth and reactivity to the BANA test of samples of subgingival microbiota from patients with periodontitis. *Braz Dent J*. 1998;9:77–84.
- Grossi SG, Dunford RG, Ho A, Koch G, Machtei EE, Genco RJ. Sources of error for periodontal probing measurements. *J Periodontol Res*. 1996;31:330–6.
- Guentsch A, Puklo M, Preshaw PM, Glockmann E, Pfister W, Potempa J, Eick S. Neutrophils in chronic and aggressive periodontitis in interaction with *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans*. *J Periodontol Res*. 2009;44:368–77.
- Guentsch A, Kramesberger M, Sroka A, Pfister W, Potempa J, Eick S. Comparison of gingival crevicular fluid sampling methods in patients with severe chronic periodontitis. *J Periodontol*. 2011;82:1051–60.
- Gürsoy UK, Könönen E, Uitto VJ, Pussinen PJ, Hyvärinen K, Suominen-Taipale L, Knuuttila M. Salivary interleukin-1b concentration and the presence of multiple pathogens in periodontitis. *J Clin Periodontol*. 2009;36:922–7.
- Gustafsson A. Methodological considerations in gingival crevicular fluid sampling with paper strips: poor recovery of uncomplexed elastase. *J Clin Periodontol*. 1996;23:432–6.
- Haffajee AD, Socransky SS, Goodson JM. Clinical parameters as predictors of destructive periodontal disease activity. *J Clin Periodontol*. 1983;10:257–65.
- Haffajee AD, Socransky SS, Goodson JM. Subgingival temperature (a). Relation to baseline clinical parameters. *J Clin Periodontol*. 1992a;19:401–8.
- Haffajee AD, Socransky SS, Goodson JM. Subgingival temperature (b). Relation to future periodontal attachment loss. *J Clin Periodontol*. 1992b;19:409–16.
- Haffajee AD, Socransky SS, Smith C, Dibart S, Goodson JM. Subgingival temperature (c). Relation to microbial counts. *J Clin Periodontol*. 1992c;19:417–22.
- Haffajee AD, Yaskell T, Torresyap G, Teles R, Socransky SS. Comparison between polymerase chain reaction-based and checkerboard DNA hybridization techniques for microbial assessment of subgingival plaque samples. *J Clin Periodontol*. 2009;36:642–9.
- Haigh BJ, Stewart KW, Whelan JR, Barnett MP, Smolenski GA, Wheeler TT. Alterations in the salivary proteome associated with periodontitis. *J Clin Periodontol*. 2010;37:241–7.
- Hanioka T, Shizukuishi S, Tsunemitsu A, Joh S, Sugiyama K, Sato N. Haemoglobin concentration and oxygen saturation in dog gingiva with experimentally induced periodontitis. *Arch Oral Biol*. 1989;34:657–63.
- Hanioka T, Shizukuishi S, Tsunemitsu A. Hemoglobin concentration and oxygen saturation of clinically healthy and inflamed gingiva in human subjects. *J Periodontol Res*. 1990;25:93–8.
- Hanioka T, Tanaka M, Ojima M, Takaya K, Matsumori Y, Shizukuishi S. Oxygen sufficiency in the gingiva of smokers and non-smokers with periodontal disease. *J Periodontol*. 2000a;71:1846–51.
- Hanioka T, Tanaka M, Takaya K, Matsumori Y, Shizukuishi S. Pocket oxygen tension in smokers and non-smokers with periodontal disease. *J Periodontol*. 2000b;71:550–4.

- Hannig C, Hannig M, Rehmer O, Braun G, Hellwig E, Al-Ahmad A. Fluorescence microscopic visualization and quantification of initial bacterial colonization on enamel in situ. *Arch Oral Biol*. 2007;52:1048–56.
- Hannig C, Follo M, Hellwig E, Al-Ahmad A. Visualization of adherent micro-organisms using different techniques. *J Med Microbiol*. 2010;59:1–7.
- Harase Y, Araki K, Okano T. Diagnostic ability of extraoral tuned aperture computed tomography (TACT) for impacted third molars. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2005;100:84–91.
- Harase Y, Araki K, Okano T. Accuracy of extraoral tuned aperture computed tomography (TACT) for proximal caries detection. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2006;101:791–6.
- Hartroth B, Seyfahrt I, Conrads G. Sampling of periodontal pathogens by paper points: evaluation of basic parameters. *Oral Microbiol Immunol*. 1999;14:326–30.
- Hatipoğlu H, Yamalik N, Berberoğlu A, Eratalay K. Impact of the distinct sampling area on volumetric features of gingival crevicular fluid. *J Periodontol*. 2007;78:705–15.
- Hausmann E. Radiographic and digital imaging in periodontal practice. *J Periodontol*. 2000;71:497–503.
- Hedin CA, Ronquist G, Forsberg O. Cyclic nucleotide content in gingival tissue of smokers and non-smokers. *J Periodontal Res*. 1981;16:337–43.
- Hefti AF. Periodontal probing. *Crit Rev Oral Biol Med*. 1997;8:336–56.
- Heitz-Mayfield LJ. Disease progression: identification of high-risk groups and individuals for periodontitis. *J Clin Periodontol*. 2005;32 Suppl 6:196–209.
- Hemmings KW, Griffiths GS, Bulman JS. Detection of neutral protease (Periocheck) and BANA hydrolase (Perioscan) compared with traditional clinical methods of diagnosis and monitoring of chronic inflammatory periodontal disease. *J Clin Periodontol*. 1997;24:110–4.
- Hernández M, Gamonal J, Tervahartiala T, Mäntylä P, Rivera O, Dezerega A, Dutzan N, Sorsa T. Associations between matrix metalloproteinase-8 and -14 and myeloperoxidase in gingival crevicular fluid from subjects with progressive chronic periodontitis: a longitudinal study. *J Periodontol*. 2010;81:1644–52.
- Herr AE, Hatch AV, Giannobile WV, Throckmorton DJ, Tran HM, Brennan JS, Singh AK. Integrated microfluidic platform for oral diagnostics. *Ann N Y Acad Sci*. 2007a;1098:362–74.
- Herr AE, Hatch AV, Throckmorton DJ, Tran HM, Brennan JS, Giannobile WV, Singh AK. Microfluidic immunoassays as rapid saliva-based clinical diagnostics. *Proc Natl Acad Sci USA*. 2007b;104:5268–73.
- Holthuis AF, Gelskey SC, Chebib FS. The relationship between gingival tissue temperatures and various indicators of gingival inflammation. *J Periodontol*. 1981;52:187–9.
- Holzhausen M, Cortelli JR, da Silva VA, Franco GC, Cortelli SC, Vergnolle N. Protease-activated receptor-2 (PAR(2)) in human periodontitis. *J Dent Res*. 2010;89:948–53.
- Horton RA, Ludlow JB, Webber RL, Gates W, Nason Jr RH, Reboussin D. Detection of peri-implant bone changes with axial tomosynthesis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 1996;81:124–9.
- Huang Y, Umeda M, Takeuchi Y, Ishizuka M, Yano-Higuchi K, Ishikawa I. Distribution of *Bacteroides forsythus* genotypes in a Japanese periodontitis population. *Oral Microbiol Immunol*. 2003;18:208–14.
- Huyghe A, Francois P, Charbonnier Y, Tangomo-Bento M, Bonetti EJ, Paster BJ, Bolivar I, Baratti-Mayer D, Pittet D, Schrenzel J, Geneva Study Group on Noma (GESNOMA). Novel microarray design strategy to study complex bacterial communities. *Appl Environ Microbiol*. 2008;74:1876–85.
- Ikezawa-Suzuki I, Shimada Y, Tai H, Komatsu Y, Tanaka A, Yoshie H. Effects of treatment on soluble tumour necrosis factor receptor type 1 and 2 in chronic periodontitis. *J Clin Periodontol*. 2008;35:961–8.
- Ishihata K, Wakabayashi N, Wadachi J, Akizuki T, Izumi Y, Takakuda K, Igarashi Y. Reproducibility of pocket depth measurement by experimental periodontal probe incorporating optical fiber sensor. *J Periodontol*. 2012;83(2):222–7. Epub 2011 May 16.
- Ishisaka A, Ansai T, Soh I, Inenaga K, Awano S, Yoshida A, Hamasaki T, Sonoki K, Takata Y, Nishihara T, Takehara T. Association of cortisol and dehydroepiandrosterone sulphate levels in serum with periodontal status in older Japanese adults. *J Clin Periodontol*. 2008;35:853–61.
- Iwasaki M, Yoshihara A, Hiroto T, Ogawa H, Hanada N, Miyazaki H. Longitudinal study on the relationship between serum albumin and periodontal disease. *J Clin Periodontol*. 2008;35:291–6.
- Jeffcoat MK, Jeffcoat RL, Jens SC, Captain K. A new periodontal probe with automated cemento-enamel junction detection. *J Clin Periodontol*. 1986;13:276–80.
- Jepsen S, Springer IN, Buschmann A, Hedderich J, Açil Y. Elevated levels of collagen cross-link residues in gingival tissues and crevicular fluid of teeth with periodontal disease. *Eur J Oral Sci*. 2003;111:198–202.
- Jervøe-Storm P-M, Koltzsch M, Falk W, Dörfler A, Jepsen S. Comparison of culture and real-time PCR for detection and quantification of five putative periodontopathogenic bacteria in subgingival plaque samples. *J Clin Periodontol*. 2005;32:778–83.
- Jervøe-Storm PM, AlAhdab H, Koltzsch M, Fimmers R, Jepsen S. Quantification of periodontal pathogens by paper point sampling from the coronal and apical aspect of periodontal lesions by real-time PCR. *Clin Oral Investig*. 2010;14:533–41.
- Jervøe-Storm PM, Alahdab H, Koltzsch M, Fimmers R, Jepsen S. Comparison of curet and paper point sampling of subgingival bacteria as analyzed by real-time polymerase chain reaction. *J Periodontol*. 2007;78:909–17.
- Kamiya I, Okuda K, Hara K. Flow-cytometric identification and detection of *Porphyromonas gingivalis* by a LPS specific monoclonal antibody. *J Periodontol*. 1994;65:309–15.

- Kamma JJ, Nakou M, Gmür R, Baehni PC. Microbiological profile of early onset/aggressive periodontitis patients. *Oral Microbiol Immunol*. 2004;19:314–21.
- Kardeşler L, Buduneli N, Biyikoğlu B, Cetinkalp S, Küttükçüler N. Gingival crevicular fluid PGE2, IL-1beta, t-PA, PAI-2 levels in type 2 diabetes and relationship with periodontal disease. *Clin Biochem*. 2008;41(10–11):863–8.
- Kardeşler L, Biyikoğlu B, Cetinkalp S, Pitkala M, Sorsa T, Buduneli N. Crevicular fluid matrix metalloproteinase-8, -13, and TIMP-1 levels in type 2 diabetics. *Oral Dis*. 2010;16:476–81.
- Karpinia K, Magnusson I, Gibbs C, Yang MC. Accuracy of probing attachment levels using a CEJ probe versus traditional probes. *J Clin Periodontol*. 2004;31:173–6.
- Kasaj A, Willershausen B. Digital volume tomography for diagnostics in periodontology. *Int J Comput Dent*. 2007;10:155–68.
- Kaufman E, Lamster IB. Analysis of saliva for periodontal diagnosis – a review. *J Clin Periodontol*. 2000;27:453–65.
- Keagle JG, Garnick JJ, Searle JR, King GE, Morse PK. Gingival resistance to probing forces. I. Determination of optimal probe diameter. *J Periodontol*. 1989;60:167–71.
- Kinane DF, Darby IB, Said S, Luoto H, Sorsa T, Tikanoja S, Mäntylä P. Changes in gingival crevicular fluid matrix metalloproteinase-8 levels during periodontal treatment and maintenance. *J Periodontol Res*. 2003;38:400–4.
- Kinane DF, Preshaw PM, Loos BG, Working Group 2 of Seventh European Workshop on Periodontology. Host-response: understanding the cellular and molecular mechanisms of host-microbial interactions – consensus of the seventh European workshop on periodontology. *J Clin Periodontol*. 2011;38 Suppl 11:44–8.
- Kinney JS, Ramseier CA, Giannobile WV. Oral fluid-based biomarkers of alveolar bone loss in periodontitis. *Ann N Y Acad Sci*. 2007;1098:230–51.
- Kivelä-Rajamäki M, Maisi P, Srinivas R, Tervahartiala T, Teronen O, Husa V, Salo T, Sorsa T. Levels and molecular forms of MMP-7 (matrilysin-1) and MMP-8 (collagenase-2) in diseased human peri-implant sulcular fluid. *J Periodontol Res*. 2003a;38:583–90.
- Kivelä-Rajamäki MJ, Teronen OP, Maisi P, Husa V, Tervahartiala TI, Piriälä EM, Salo TA, Mellanen L, Sorsa TA. Laminin-5 gamma2-chain and collagenase-2 (MMP-8) in human peri-implant sulcular fluid. *Clin Oral Implants Res*. 2003b;14:158–65.
- Kornman KS. Reaction: sampling of micro-organisms associated with periodontal disease. *Oral Microbiol Immunol*. 1986;1:21–2.
- Koss MA, Castro CE, Salúm KM, López ME. Changes in saliva protein composition in patients with periodontal disease. *Acta Odontol Latinoam*. 2009;22:105–12.
- Krigr DM, Kaltschmitt J, Krieger JK, Eickholz P. Two subgingival plaque-sampling strategies used with RNA probes. *J Periodontol*. 2007;78:72–8.
- Kuboniwa M, Inaba H, Amano A. Genotyping to distinguish microbial pathogenicity in periodontitis. *Periodontol*. 2000;54:136–59.
- Kung RT, Ochs B, Goodson JM. Temperature as a periodontal diagnostic. *J Clin Periodontol*. 1990;17:557–63.
- Kurtiş B, Develioğlu H, Taner IL, Baloş K, Tekin IO. IL-6 levels in gingival crevicular fluid (GCF) from patients with non-insulin dependent diabetes mellitus (NIDDM), adult periodontitis and healthy subjects. *J Oral Sci*. 1999;41:163–7.
- Kwack L, Smith JK, Castillo M, Ewend MG, Cush S, Hensing T, et al. Clinical applications of proton MR spectroscopy in oncology. *Technol Cancer Res Treat*. 2002;1:17–28.
- Lakio L, Kuula H, Dogan B, Asikainen S. *Actinobacillus actinomycetemcomitans* proportion of subgingival bacterial flora in relation to its clonal type. *Eur J Oral Sci*. 2002;110:212–7.
- Lamster IB. Evaluation of components of gingival crevicular fluid as diagnostic tests. *Ann Periodontol*. 1997;2:123–37.
- Lamster IB, Ahlo JK. Analysis of gingival crevicular fluid as applied to the diagnosis of oral and systemic diseases. *Ann N Y Acad Sci*. 2007;1098:216–29.
- Lamster IB, Oshrain RL, Gordon JM. Enzyme activity in human gingival crevicular fluid: considerations in data reporting based on analysis of individual crevicular sites. *J Clin Periodontol*. 1986;13:799–804.
- Lamster IB, Oshrain RL, Fiorello LA, Celenti RS, Gordon JM. A comparison of 4 methods of data presentation for lysosomal enzyme activity in gingival crevicular fluid. *J Clin Periodontol*. 1988;15:347–52.
- Lamster IB, Holmes LG, Gross KB, Oshrain RL, Cohen DW, Rose LF, Peters LM, Pope MR. The relationship of beta-glucuronidase activity in crevicular fluid to clinical parameters of periodontal disease. Findings from a multicenter study. *J Clin Periodontol*. 1994;21:118–27.
- Lamster IB, Kaufman E, Grbic JT, Winston LJ, Singer RE. Beta-glucuronidase activity in saliva: relationship to clinical periodontal parameters. *J Periodontol*. 2003;74:353–9.
- Lang NP, Adler R, Joss A, Nyman S. Absence of bleeding on probing. An indicator of periodontal stability. *J Clin Periodontol*. 1990;17:714–21.
- Larsen C, Barendregt DS, Slot DE, Van der Velden V, Van der Weijden F. Probing pressure, a highly undervalued unit of measure in periodontal probing: a systematic review on its effect on probing pocket depth. *J Clin Periodontol*. 2009;36:315–22.
- Lau L, Sanz M, Herrera D, Morillo JM, Martín C, Silva A. Quantitative real-time polymerase chain reaction versus culture: a comparison between two methods for the detection and quantification of *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis* and *Tannerella forsythensis* in subgingival plaque samples. *J Clin Periodontol*. 2004;31:1061–9.
- Lee SS, Kim HK, Choi SC, Lee JI. Granulocytic sarcoma occurring in the maxillary gingiva demonstrated by magnetic resonance imaging. *Oral Surg*

- Oral Med Oral Pathol Oral Radiol Endod. 2001;92: 689–93.
- Lee HM, Cincio SG, Tüter G, Ryan ME, Komaroff E, Golub LM. Subantimicrobial dose doxycycline efficacy as a matrix metalloproteinase inhibitor in chronic periodontitis patients is enhanced when combined with a non-steroidal anti-inflammatory drug. *J Periodontol*. 2004;75:453–63.
- Lee JM, Garon E, Wong DT. Salivary diagnostics. *Orthod Craniofac Res*. 2009;12:206–11.
- Leroy R, Eaton KA, Savage A. Methodological issues in epidemiological studies of periodontitis – how can it be improved? *BMC Oral Health*. 2010;10:8.
- Li Y, Denny P, Ho CM, Montemagno C, Shi W, Qi F, Wu B, Wolinsky L, Wong DT. The Oral Fluid MEMS/ NEMS Chip (OFMNC): diagnostic and translational applications. *Adv Dent Res*. 2005;18:3–5.
- Lima KC, Coelho LT, Pinheiro IV, Rôças IN, Siqueira Jr JF. Microbiota of dentinal caries as assessed by reverse-capture checkerboard analysis. *Caries Res*. 2011;45:21–30.
- Limrachtamorn S, Edge MJ, Gettleman L, Scheetz JP, Farman AG. Array geometry for assessment of mandibular implant position using tuned aperture computed tomography (TACT). *Dentomaxillofac Radiol*. 2004;33:25–31.
- Lin SJ, Chen YL, Kuo MY, Li CL, Lu HK. Measurement of gp130 cytokines oncostatin M and IL-6 in gingival crevicular fluid of patients with chronic periodontitis. *Cytokine*. 2005;30:160–7.
- Listgarten MA, Levin S. Positive correlation between the proportions of subgingival spirochetes and motile bacteria and susceptibility of human subjects to periodontal deterioration. *J Clin Periodontol*. 1981;8(2):122–38.
- Liu KZ, Xiang XM, Man A, Sowa MG, Cholakias A, Ghiabi E, Singer DL, Scott DA. In vivo determination of multiple indices of periodontal inflammation by optical spectroscopy. *J Periodontal Res*. 2009;44: 117–24.
- Liu J, Wu Y, Ding Y, Meng S, Ge S, Deng H. Evaluation of serum levels of C-reactive protein and lipid profiles in patients with chronic periodontitis and/or coronary heart disease in an ethnic Han population. *Quintessence Int*. 2010;41:239–47.
- Lockhart PB, Kim S, Lund NL. Magnetic resonance imaging of human teeth. *J Endod*. 1992;18:237–44.
- Loomer PM. Microbiological diagnostic testing in the treatment of periodontal diseases. *Periodontol* 2000. 2004;34:49–56.
- Loos BG, Dyer DW. Restriction fragment length polymorphism analysis of the fimbriin locus, fimA, of *Porphyromonas gingivalis*. *J Dent Res*. 1992;71: 1173–81.
- Loos BG, Tjoa S. Host-derived diagnostic markers for periodontitis: do they exist in gingival crevice fluid? *Periodontol* 2000. 2005;39:53–72.
- Loos BG, Mayrand D, Genco RJ, Dickinson DP. Genetic heterogeneity of *Porphyromonas* (Bacteroides) *gingivalis* by genomic DNA fingerprinting. *J Dent Res*. 1990;69:1488–93.
- Loos BG, Dyer DW, Whittam TS, Selander RK. Genetic structure of populations of *Porphyromonas gingivalis* associated with periodontitis and other oral infections. *Infect Immun*. 1993;61:204–12.
- Löst C, Irion KM, Nüssle W. Ultrasonic B-scans of the facial/oral periodontium in pigs. *J Clin Periodontol*. 1989;16:534–8.
- Loubele M, van Assche N, Carpentier K, et al. Comparative localized linear accuracy of small-field cone-beam CT and multislice CT for alveolar bone measurements. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2008;105:512–8.
- Lynch JE, Hinders MK. Ultrasonic device for measuring periodontal attachment levels. *Rev Sci Instrum*. 2002;73:2686–93.
- Lynch JE, Hinders MK, McCombs GB. Clinical comparison of an ultrasonographic periodontal probe to manual and controlled-force probing. *Measurement*. 2006;39:429–39.
- Magnusson I, Fuller WW, Heins PJ, Rau CF, Gibbs CH, Marks RG, et al. Correlation between electronic and visual readings of pocket depths with a newly developed constant force probe. *J Clin Periodontol*. 1988;15:180–4.
- Mahmoud AM, Ngan P, Crout R, Mukdadi OM. High-resolution 3D ultrasound jawbone surface imaging for diagnosis of periodontal bony defects: an in vitro study. *Ann Biomed Eng*. 2010;38:3409–22.
- Maiden MF, Tanner AC, Macuch PJ, Murray L, Kent Jr RL. Subgingival temperature and microbiota in initial periodontitis. *J Clin Periodontol*. 1998;25:786–93.
- Mandel ID. The diagnostic uses of saliva. *J Oral Pathol Med*. 1990;19:119–25.
- Mäntylä P, Stenman M, Kinane DF, Tikanoja S, Luoto H, Salo T, Sorsa T. Gingival crevicular fluid collagenase-2 (MMP-8) test stick for chair-side monitoring of periodontitis. *J Periodontal Res*. 2003;38:436–9.
- Mäntylä P, Stenman M, Kinane D, Salo T, Suomalainen K, Tikanoja S, Sorsa T. Monitoring periodontal disease status in smokers and nonsmokers using a gingival crevicular fluid matrix metalloproteinase-8-specific chair-side test. *J Periodontal Res*. 2006;41:503–12.
- Marcaccini AM, Meschiari CA, Sorgi CA, Saraiva MC, de Souza AM, Faccioli LH, Tanus-Santos JE, Novaes AB, Gerlach RF. Circulating interleukin-6 and high sensitivity C-reactive protein decrease after periodontal therapy in otherwise healthy subjects. *J Periodontol*. 2009;80:594–602.
- McCulloch CA, Birek P, Hardy V. Comparison of gingival attachment level measurements with an automated periodontal probe and a pressure-sensitive probe. *J Periodontal Res*. 1987;22:348–52.
- Mellado JR, Freedman AL, Salkin LM, Stein MD, Schneider DB, Cutler RH. The clinical relevance of microbiologic testing: a comparative analysis of microbiologic samples secured from the same sites and cultured in two independent laboratories. *Int J Periodontics Restorative Dent*. 2001;21:232–9.
- Menard C, Mouton C. Randomly amplified polymorphic DNA analysis confirms the biotyping scheme of *Porphyromonas gingivalis*. *Res Microbiol*. 1993;144:445–55.

- Menard C, Mouton C. Clonal diversity of the taxon *Porphyromonas gingivalis* assessed by random amplified polymorphic DNA fingerprinting. *Infect Immun*. 1995;63:2522–31.
- Menard C, Brousseau R, Mouton C. Application of polymerase chain reaction with arbitrary primer (AP-PCR) to strain identification of *Porphyromonas* (Bacteroides) *gingivalis*. *FEMS Microbiol Lett*. 1992;74:163–8.
- Menard C, Gosselin P, Duhaime JF, Mouton C. Polymerase chain reaction using arbitrary primer for the design and construction of a DNA probe specific for *Porphyromonas gingivalis*. *Res Microbiol*. 1994;145:595–602.
- Mengel R, Candir M, Shiratori K, et al. Digital volume tomography in the diagnosis periodontal defects: an in vitro study on native pig and human mandibles. *J Periodontol*. 2005;76:665–73.
- Meurman JH, Wahlfors J, Korhonen A, et al. Identification of *Bacteroides forsythus* in subgingival dental plaque with the aid of a rapid PCR method. *J Dent Res*. 1997;76:1376–80.
- Miller CS, Foley JD, Bailey AL, Campell CL, Humphries RL, Christodoulides N, Floriano PN, Simmons G, Bhagwandin B, Jacobson JW, Redding SW, Ebersole JL, McDevitt JT. Current developments in salivary diagnostics. *Biomark Med*. 2010;4:171–89.
- Misch KA, Yi ES, Sarment DP. Accuracy of cone beam computed tomography for periodontal defect measurements. *J Periodontol*. 2006;77:1261–6.
- Mjor IA, Webber RL, Horton RA. Computerized tomosynthetic radiography in operative dentistry. *Quintessence Int*. 1997;28:99–103.
- Mogi M, Ootogoto J. Expression of cathepsin-K in gingival crevicular fluid of patients with periodontitis. *Arch Oral Biol*. 2007;52:894–8.
- Mogi M, Ootogoto J, Ota N, Togari A. Differential expression of RANKL and osteoprotegerin in gingival crevicular fluid of patients with periodontitis. *J Dent Res*. 2004;83:166–9.
- Mombelli A, Graf H. Depth-force-patterns in periodontal probing. *J Clin Periodontol*. 1986;13:126–30.
- Mombelli A, Graf H. Depth-force-patterns in periodontal probing. *J Clin Periodontol*. 1986;13:126–30.
- Mol A. Imaging methods in periodontology. *Periodontol*. 2000;34:34–48.
- Mol A, Balasundaram A. In vitro cone beam computed tomography imaging of periodontal bone. *Dentomaxillofac Radiol*. 2008;37:1–6.
- Mombelli A, Minder CE, Gusberti FA, Lang NP. Reproducibility of microscopic and cultural data in repeated subgingival plaque samples. *J Clin Periodontol*. 1989;16:434–42.
- Mombelli A, McNabb H, Lang NP. Black-pigmenting Gram-negative bacteria in periodontal disease. I. Topographic distribution in the human dentition. *J Periodontol Res*. 1991a;26:301–7.
- Mombelli A, McNabb H, Lang NP. Black-pigmenting Gram-negative bacteria in periodontal disease. II. Screening strategies for detection of *P. gingivalis*. *J Periodontol Res*. 1991b;26:308–13.
- Mombelli A, Muhle T, Frigg R. Depth-force patterns of periodontal probing. Attachment-gain in relation to probing force. *J Clin Periodontol*. 1992;19:295–300.
- Mombelli A, Gmür R, Gobbi C, Lang NP. *Actinobacillus actinomycetemcomitans* in adult periodontitis. I. Topographic distribution before and after treatment. *J Periodontol*. 1994a;65:820–6.
- Mombelli A, Gmür R, Gobbi C, Lang NP. *Actinobacillus actinomycetemcomitans* in adult periodontitis. II. Characterization of isolated strains and effect of mechanical periodontal treatment. *J Periodontol*. 1994b;65:827–34.
- Morikawa M, Chiba T, Tomii N, Sato S, Takahashi Y, Konishi K, Numabe Y, Iwata K, Imai K. Comparative analysis of putative periodontopathic bacteria by multiplex polymerase chain reaction. *J Periodontol Res*. 2008;43:268–74.
- Morillo JM, Lau L, Sanz M, Herrera D, Martín C, Silva A. Quantitative real-time polymerase chain reaction based on single copy gene sequence for detection of periodontal pathogens. *J Clin Periodontol*. 2004;31:1054–60.
- Morozumi T, Kubota T, Sato T, Okuda K, Yoshie H. Smoking cessation increases gingival blood flow and gingival crevicular fluid. *J Clin Periodontol*. 2004;31:267–72.
- Mousques T, Listgarten MA, Stoller NH. Effect of sampling on the composition of the human subgingival microbial flora. *J Periodontol Res*. 1980;15:137–43.
- Müller HP, Heinecke A. The influence of gingival dimensions on bleeding upon probing in young adults with plaque-induced gingivitis. *Clin Oral Investig*. 2002;6:69–74.
- Müller HP, Heinecke A, Eger T. Site-specific association between supragingival plaque and bleeding upon probing in young adults. *Clin Oral Investig*. 2000;4:212–8.
- Müller HP, Stadermann S, Heinecke A. Longitudinal association between plaque and gingival bleeding in smokers and non-smokers. *J Clin Periodontol*. 2002;29:287–94.
- Munjal S, Miethe P, Netuschil L, Struck F, Maier K, Bauermeister C. Immunoassay-based diagnostic point-of-care technology for oral specimen. *Ann N Y Acad Sci*. 2007a;1098:486–9.
- Munjal SK, Prescher N, Struck F, Sorsa T, Maier K, Netuschil L. Evaluation of immunoassay-based MMP-8 detection in gingival crevicular fluid on a point-of-care platform. *Ann N Y Acad Sci*. 2007b;1098:490–2.
- Na J, Lee BH, Baek JH, Choi ES. Optical approach for monitoring the periodontal ligament changes induced by orthodontic forces around maxillary anterior teeth of white rats. *Med Biol Eng Comput*. 2008;46:597–603.
- Nair MK, Bezik J. Tuned-aperture computed tomography for detection of induced mid-buccal/lingual alveolar bone defects. *J Periodontol*. 2006;77:1833–8.
- Nair MK, Tyndall DA, Ludlow JB, May K. Tuned aperture computed tomography and detection of recurrent caries. *Caries Res*. 1998;32:23–30.
- Nair MK, Seyedain A, Agarwal S, Webber RL, Nair UP, Piesco NP, Mooney MP, Grondahl HG. Tuned aperture computed tomography to evaluate osseous healing. *J Dent Res*. 2001;80:1621–4.
- Nair MK, Nair UP, Gröndahl HG, Webber RL. Accuracy of tuned aperture computed tomography in the diagnosis of radicular fractures in non-restored maxillary

- anterior teeth--an in vitro study. *Dentomaxillofac Radiol.* 2002;31:299-304.
- Nair MK, Gröndahl HG, Webber RL, Nair UP, Wallace JA. Effect of iterative restoration on the detection of artificially induced vertical radicular fractures by tuned aperture computed tomography. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2003a;96:118-25.
- Nair P, Sutherland G, Palmer RM, Wison RF, Scott DA. Gingival bleeding on probing increases after quitting smoking. *J Clin Periodontol.* 2003b;30:435-7.
- Nair MK, Nair UP, Seyedain A, Gassner R, Piesco N, Mooney M, Ganta S, Agarwal S. Correlation of tuned aperture computed tomography with conventional computed tomography for evaluation of osseous healing in calvarial defects. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2007;103:267-73.
- Nakagawa M, Kurihara H, Nishimura F, et al. Immunological, genetic, and microbiological study of family members manifesting early-onset periodontitis. *J Periodontol.* 1996;67:254-63.
- Nair MK, Nair UP, Gröndahl HG, Webber RL. Accuracy of tuned aperture computed tomography in the diagnosis of radicular fractures in non-restored maxillary anterior teeth--an in vitro study. *Dentomaxillofac Radiol.* 2002;31:299-304.
- Nakajima T, Honda T, Domon H, Okui T, Kajita K, Ito H, Takahashi N, Mackawa T, Tabeta K, Yamazaki K. Periodontitis-associated up-regulation of systemic inflammatory mediator level may increase the risk of coronary heart disease. *J Periodontol Res.* 2010;45: 116-22.
- Nakamura-Minami M, Furuichi Y, Ishikawa K, Mitsuzono-Tofuku Y, Izumi Y. Changes of alpha1-protease inhibitor and secretory leukocyte protease inhibitor levels in gingival crevicular fluid before and after non-surgical periodontal treatment. *Oral Dis.* 2003;9:249-54.
- Navazesh M. Methods for collecting saliva. *Ann N Y Acad Sci.* 1993;694:72-7.
- Neto JB, Filho GR, Tramontina VA, Sallum EA, Nociti Jr FH, Sallum AW. Millimeter marks and probe tip diameter standardisation from commercially available periodontal probes. A comparative study. *J Int Acad Periodontol.* 2001;3:57-60.
- Nibali L, D' Aiuto F, Griffiths G, Patel K, Suvan J, Tonetti MS. Severe periodontitis is associated with systemic inflammation and a dysmetabolic status: a case-control study. *J Clin Periodontol.* 2007;34:931-7.
- Nicu EA, Laine ML, Morre SA, Van der Velden U, Loos BG. Soluble CD14 in periodontitis. *Innate Immun.* 2009;15:121-8.
- Niederman R, Kent R. Use of subgingival temperature in periodontal clinical trials. Assessment of accuracy and reliability. *J Periodontol Res.* 1993;28(6 Pt 2):540-2.
- Nisengard RC, Newman MG, Sanz M. Host response: basic concepts. In: Carranza FA, Newman MG, editors. *Clinical periodontology*. London: WB Saunders Company; 2006.
- Nogueira-Filho G, Xiang XM, Shibli JA, Duarte PM, Sowa MG, Ferrari DS, Onuma T, de Cardoso LA, Liu KZ. On site noninvasive assessment of peri-implant inflammation by optical spectroscopy. *J Periodontol Res.* 2011;46:382-8.
- Nomura Y, Tamaki Y, Tanaka T, Arakawa H, Tsurumoto A, Kirimura K, Sato T, Hanada N, Kamoi K. Screening of periodontitis with salivary enzyme tests. *J Oral Sci.* 2006;48:177-83.
- Nonnenmacher C, Dalpke A, Rochon J, Flores-de-Jacoby L, Mutters R, Heeg K. Real-time polymerase chain reaction for detection and quantification of bacteria in periodontal patients. *J Periodontol.* 2005;76:1542-9.
- Noujeim M, Nummikoski P, Langlais R. Evaluation of high-resolution cone-beam computed tomography in the detection of simulated interradicular bone lesions. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2007;103:e52.
- Offenbacher S, Beck JD, Moss K, Mendoza L, Paquette DW, Barrow DA, Couper DJ, Stewart DD, Falkner KL, Graham SP, Grossi S, Gunsolley JC, Madden T, Maupome G, Trevisan M, Van Dyke TE, Genco RJ. Results from the Periodontitis and Vascular Events (PAVE) Study: a pilot multicentered, randomized, controlled trial to study effects of periodontal therapy in a secondary prevention model of cardiovascular disease. *J Periodontol.* 2009;80:190-201.
- Okada M, Hayashi F, Nagasaka N. Detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis* in dental plaque samples from children 2 to 12 years of age. *J Clin Periodontol.* 2000;27:763-8.
- Okada M, Hayashi F, Nagasaka N. PCR detection of 5 putative periodontal pathogens in dental plaque samples from children 2 to 12 years of age. *J Clin Periodontol.* 2001;28:576-82.
- Oringer RJ, Al-Shammari KF, Aldredge WA, Iacono VJ, Eber RM, Wang HL, Berwald B, Nejat R, Giannobile WV. Effect of locally delivered minocycline microspheres on markers of bone resorption. *J Periodontol.* 2002;73:835-42.
- Osborn J, Stoltenberg J, Huso B, Aeppli D, Pihlstrom B. Comparison of measurement variability using a standard and constant force periodontal probe. *J Periodontol.* 1990;61:497-503.
- Otis LL, Everett MJ, Sathyam US, Colston Jr BW. Optical coherence tomography: a new imaging technology for dentistry. *J Am Dent Assoc.* 2000;131:511-4.
- Ozkavaf A, Aras H, Huri CB, Mottaghian-Dini F, Tözüm TF, Etikan I, Yamalik N, Caglayan F. Relationship between the quantity of gingival crevicular fluid and clinical periodontal status. *J Oral Sci.* 2000;42:231-8.
- Ozkavaf A, Aras H, Huri CB, et al. Analysis of factors that may affect the enzymatic profile of gingival crevicular fluid: sampling technique, sequential sampling and mode of data presentation. *J Oral Sci.* 2001;43:41-8.
- Ozmeric N. Advances in periodontal disease markers. *Clin Chim Acta.* 2004;343(1-2):1-16.
- Özmeric N, Elgün S, Uraz A. Salivary arginase in patients with adult periodontitis. *Clin Oral Investig.* 2000;4:21-4.
- Palanisamy V, Sharma S, Deshpande A, Zhou H, Gimzewski J, Wong DT. Nanostructural and transcriptomic analyses of human saliva derived exosomes. *PLoS One.* 2010;5:e8577.
- Palys MD, Haffajee AD, Socransky SS, Giannobile WV. Relationship between C-telopeptide pyridinoline

- cross-links (ICTP) and putative periodontal pathogens in periodontitis. *J Clin Periodontol.* 1998;25(11 Pt 1): 865–71.
- Papapanou PN, Madianos PN, Dahlén G, Sandros J. “Checkerboard” versus culture: a comparison between two methods for identification of subgingival microbiota. *Eur J Oral Sci.* 1997;105:389–96.
- Parr GR, Bustos-Valdes SE. A modified segregator for collection of human submandibular and sublingual saliva. *Arch Oral Biol.* 1984;29:69–71.
- Passoja A, Ylipalosaari M, Tervonen T, Raunio T, Knuuttila M. Matrix metalloproteinase-8 concentration in shallow crevices associated with the extent of periodontal disease. *J Clin Periodontol.* 2008;35:1027–31.
- Paster BJ, Dewhirst FE. Molecular microbial diagnosis. *Periodontol* 2000. 2009;51:38–44.
- Paster BJ, Bartoszyk IM, Dewhirst FE. Identification of oral streptococci using PCR-based, reverse-capture, checkerboard hybridization. *Methods Cell Sci.* 1998;20:223–31.
- Paster BJ, Russell MK, Alpagot T, Lee AM, Boches SK, Galvin JL, Dewhirst FE. Bacterial diversity in necrotizing ulcerative periodontitis in HIV-positive subjects. *Ann Periodontol.* 2002;7:8–16.
- Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. *Periodontol* 2000. 2006;42: 80–7.
- Perez-Chaparro PJ, Gracieux P, Lafaurie GI, Donnio PY, Bonnaure-Mallet M. Genotypic characterization of *Porphyromonas gingivalis* isolated from subgingival plaque and blood sample in positive bacteremia subjects with periodontitis. *J Clin Periodontol.* 2008;35:748–53.
- Perez-Chaparro PJ, Rouillon A, Minet J, Lafaurie GI, Bonnaure-Mallet M. fimA genotypes and PFGE profile patterns in *Porphyromonas gingivalis* isolates from subjects with periodontitis. *Oral Microbiol Immunol.* 2009;24:423–6.
- Perinetti G, Paolantonio M, Femminella B, Serra E, Spoto G. Gingival crevicular fluid alkaline phosphatase activity reflects periodontal healing/recurrent inflammation phases in chronic periodontitis patients. *J Periodontol.* 2008;79:1200–7.
- Perry DA, Taggart EJ, Leung A, Newburn E. Comparison of a conventional probe with electronic and manual pressure-regulated probes. *J Periodontol.* 1994;65: 908–13.
- Persson GR, Page RC. Effect of sampling time and repetition on gingival crevicular fluid and aspartate aminotransferase. *J Periodontal Res.* 1990;25:236–42.
- Persson L, Bergström J, Gustafsson A, Asman B. Tobacco smoking and gingival neutrophil activity in young adults. *J Clin Periodontol.* 1999;26:9–13.
- Persson GR, Weibel M, Hirschi R, Katsoulis J. Similarities in the subgingival microbiota assessed by a curet sampling method at sites with chronic periodontitis. *J Periodontol.* 2008;79:2290–6.
- Pihlstrom BL. Measurement of attachment level in clinical trials: probing methods. *J Periodontol.* 1992;63(12 Suppl):1072–7.
- Pollanen MT, Salonen JI, Uitto VJ. Structure and function of the tooth – epithelial interface in health and disease. *Periodontol* 2000. 2003;31:12–31.
- Polson AM, Caton IB, Yeaple RN, et al. Histological determination of probe tip penetration into gingival sulcus of humans using an electronic pressure sensitive probe. *J Clin Periodontol.* 1980;7:479–88.
- Pozo P, Valenzuela MA, Melej C, Zaldívar M, Puente J, Martínez B, Gamonal J. Longitudinal analysis of metalloproteinases, tissue inhibitors of metalloproteinases and clinical parameters in gingival crevicular fluid from periodontitis-affected patients. *J Periodontal Res.* 2005;40:199–207.
- Pradeep AR, Daisy H, Hadge P, Garg G, Thorat M. Correlation of gingival crevicular fluid interleukin-18 and monocyte chemoattractant protein-1 levels in periodontal health and disease. *J Periodontol.* 2009a;80:1454–61.
- Pradeep AR, Daisy H, Hadge P. Serum levels of monocyte chemoattractant protein-1 in periodontal health and disease. *Cytokine.* 2009b;47:77–81.
- Pradeep AR, Manojkumar ST, Garima G, Raju A. Serum levels of oncostatin M (a gp 130 cytokine): an inflammatory biomarker in periodontal disease. *Biomarkers.* 2010;15:277–82.
- Prescher N, Maier K, Munjal SK, Sorsa T, Bauermeister CD, Struck F, Netuschil L. Rapid quantitative chair-side test for active MMP-8 in gingival crevicular fluid: first clinical data. *Ann N Y Acad Sci.* 2007;1098: 493–5.
- Preshaw PM, Kelly PJ, Heasman PA. Quadratic calibration curves for the Periotron 6000. *J Periodontal Res.* 1996;31:441–3.
- Preshaw PM, Kupp L, Hefti AF, Mariotti A. Measurement of clinical attachment levels using a constant force periodontal probe modified to detect the cement-enamel junction. *J Clin Periodontol.* 1999;26:434–40.
- Preus HR, Zambon JJ, Dunford RG, Genco RJ. The distribution and transmission of *Actinobacillus actinomycetemcomitans* in families with established adult periodontitis. *J Periodontol.* 1994;65:2–7.
- Preza D, Olsen I, Willumsen T, Boches SK, Cotton SL, Grinde B, Paster BJ. Microarray analysis of the microflora of root caries in elderly. *Eur J Clin Microbiol Infect Dis.* 2008;46:2015–21.
- Preza D, Olsen I, Willumsen T, Boches SK, Cotton SL, Grinde B, Paster BJ. Microarray analysis of the microflora of root caries in elderly. *Eur J Clin Microbiol Infect Dis.* 2009a;28:509–17.
- Preza D, Olsen I, Willumsen T, Grinde B, Paster BJ. Diversity and site-specificity of the oral microflora in the elderly. *Eur J Clin Microbiol Infect Dis.* 2009b;28:1033–40.
- Pussinen PJ, Paju S, Mäntylä P, Sorsa T. Serum microbial and host-derived markers of periodontal diseases: a review. *Curr Med Chem.* 2007;14(22):2402–12.
- Quirynen M, Callens A, van Steenberghe D, Nys M. Clinical evaluation of a constant force electronic probe. *J Periodontol.* 1993;64:35–9.
- Ramachandra SS, Mehta DS, Sandesh N, Baliga V, Amarnath J. Periodontal probing systems: a review of

- available equipment. *Compend Contin Educ Dent*. 2011;32:71–7.
- Ramesh A, Ludlow JB, Webber RL, Tyndall DA, Paquette D. Evaluation of tuned aperture computed tomography (TACT) in the localization of simulated periodontal defects. *Dentomaxillofac Radiol*. 2001;30:319–24.
- Ramesh A, Ludlow JB, Webber RL, Tyndall DA, Paquette D. Evaluation of tuned-aperture computed tomography in the detection of simulated periodontal defects. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2002;93:341–9.
- Rammelsberg P, Jager L, Duc JM. Magnetic resonance imaging based joint space measurements in temporomandibular joints with disk displacements and in controls. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2000;90:240–8.
- Rams TE, Slots J. Comparison of two pressure-sensitive periodontal probes and a manual periodontal probe in shallow and deep pockets. *Int J Periodontics Restorative Dent*. 1993;13:520–9.
- Ramseier CA, Kinney JS, Herr AE, Braun T, Sugai JV, Shelburne CA, et al. Identification of pathogen and host-response markers correlated with periodontal disease. *J Periodontol*. 2009;80:436–46.
- Rashedi B, Tyndall DA, Ludlow JB, Chaffee NR, Guckes AD. Tuned aperture computed tomography (TACT) for cross-sectional implant site assessment in the posterior mandible. *J Prosthodont*. 2003;12:176–86.
- Raunio T, Nixdorf M, Knuuttila M, Karttunen R, Vainio O, Tervonen T. The extent of periodontal disease and the IL-6–174 genotype as determinants of serum IL-6 level. *J Clin Periodontol*. 2007;34:1025–30.
- Raunio T, Knuuttila M, Karttunen R, Vainio O, Tervonen T. Serum sCD14, polymorphism of CD14(–260) and periodontal infection. *Oral Dis*. 2009;15:484–9.
- Ready D, D'Aiuto F, Spratt DA, Suvan J, Tonetti MS, Wilson M. Disease severity associated with presence in subgingival plaque of *Porphyromonas gingivalis*, *Aggregatibacter actinomycetemcomitans*, and *Tannerella forsythia*, singly or in combination, as detected by nested multiplex PCR. *J Clin Microbiol*. 2008;46:3380–3.
- Reinhardt RA, Stoner JA, Golub LM, Lee HM, Nummikoski PV, Sorsa T, Payne JB. Association of gingival crevicular fluid biomarkers during periodontal maintenance with subsequent progressive periodontitis. *J Periodontol*. 2010;81:251–9.
- Renvert S, Wikstrom M, Helmersson M, Dahlen G, Claffey N. Comparative study of subgingival microbiological sampling techniques. *J Periodontol*. 1992;63:797–801.
- Renvert S, Lindahl C, Roos-Jansaker AM, Lessem J. Short-term effects of an anti-inflammatory treatment on clinical parameters and serum levels of C-reactive protein and proinflammatory cytokines in subjects with periodontitis. *J Periodontol*. 2009;80:892–900.
- Rescala B, Rosalem Jr W, Teles RP, Fischer RG, Haffajee AD, Socransky SS, Gustafsson A, Figueredo CM. Immunologic and microbiologic profiles of chronic and aggressive periodontitis subjects. *J Periodontol*. 2010;81:1308–16.
- Reynolds JJ, Meikle MC. Mechanisms of connective tissue matrix destruction in periodontitis. *Periodontol*. 2000. 1997;14:144–57.
- Riep B, Purucker P, Bernimoulin JP. Repeated local metronidazole-therapy as adjunct to scaling and root planing in maintenance patients. *J Clin Periodontol*. 1999;26:710–5.
- Rijnsburger MC, Delwel S, Boutaga K, Van Steenberghe TJ, Savelkoul PH, Van Winkelhoff AJ. Comparison of restriction enzyme analysis and amplified fragment length polymorphism typing of *Porphyromonas gingivalis* isolated from spouses. *Oral Microbiol Immunol*. 2007;22:369–73.
- Riviere GR, Smith KS, Willis SG, Riviere KH. Phenotypic and genotypic heterogeneity among cultivable pathogen-related oral spirochetes and *Treponema vincentii*. *J Clin Microbiol*. 1999;37:3676–80.
- Robinson PJ, Vitek RM. The relationship between gingival inflammation and resistance to probe penetration. *J Periodontal Res*. 1979;14:239–43.
- Rôças IN, Siqueira Jr JF. Root canal microbiota of teeth with chronic apical periodontitis. *J Clin Microbiol*. 2008;46:3599–606.
- Rôças IN, Siqueira Jr JF. Identification of bacteria enduring endodontic treatment procedures by a combined reverse transcriptase-polymerase chain reaction and reverse-capture checkerboard approach. *J Endod*. 2010;36:45–52.
- Rôças IN, Alves FR, Santos AL, Rosado AS, Siqueira Jr JF. Apical root canal microbiota as determined by reverse-capture checkerboard analysis of cryogenically ground root samples from teeth with apical periodontitis. *J Endod*. 2010;36:1617–21.
- Rosin M, Hanschke M, Splieth C, Kramer A. Activities of lysozyme and salivary peroxidase in unstimulated whole saliva in relation to plaque and gingivitis scores in healthy young males. *Clin Oral Investig*. 1999;3:133–7.
- Rumpf RW, Griffen AL, Leys EJ. Phylogeny of *Porphyromonas gingivalis* by ribosomal intergenic spacer region analysis. *J Clin Microbiol*. 2000;38:1807–10.
- Sahingur SE, Cohen RE. Analysis of host responses and risk for disease progression. *Periodontol*. 2000. 2004;34:57–83.
- Sakamoto M, Takeuchi Y, Umeda M, Ishikawa I, Benno Y. Rapid detection and quantification of five periodontopathic bacteria by real-time PCR. *Microbiol Immunol*. 2001;45:39–44.
- Sakamoto M, Takeuchi Y, Umeda M, Ishikawa I, Benno Y. Application of terminal RFLP analysis to characterize oral bacterial flora in saliva of healthy subjects and patients with periodontitis. *J Med Microbiol*. 2003;52:79–89.
- Sakamoto M, Huang Y, Ohnishi M, Umeda M, Ishikawa I, Benno Y. Changes in oral microbial profiles after periodontal treatment as determined by molecular analysis of 16S rRNA genes. *J Med Microbiol*. 2004;53(Pt 6):563–71.
- Sakamoto M, Umeda M, Benno Y. Molecular analysis of human oral microbiota. *J Periodontal Res*. 2005;40:277–85.

- Salkin LM, Freedman AL, Mellado JR, Stein MD, Schneider DB, Butler L. The clinical relevance of microbiologic testing. Part 2: a comparative analysis of microbiologic samples secured simultaneously from the same sites and cultured in the same laboratory. *Int J Periodontics Restorative Dent*. 2003; 23:121–7.
- Santangelo R, D’Ercole S, Graffeo R, Marchetti S, Deli G, Nacci A, Piccolomini R, Cattani P, Fadda G. Bacterial and viral DNA in periodontal disease: a study using multiplex PCR. *New Microbiol*. 2004;27:133–7.
- Sanz M, Lau L, Herrera D, Morillo JM, Silva A. Methods of detection of *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis* and *Tannerella forsythensis* in periodontal microbiology, with special emphasis on advanced molecular techniques: a review. *J Clin Periodontol*. 2004;31:1034–47.
- Sanz M, Newman MG, Quirynen M. Advanced diagnostic techniques. In: Newman MG, Takei H, Klokkevold PR, Carranza FA, editors. *Carranza’s Clinical periodontology*. 11th ed. St. Louis: Saunders; 2011. p. 579–601.
- Sato T, Kuramitsu HK. Restriction fragment-length polymorphism analysis of 16S ribosomal RNA genes amplified by polymerase chain reaction for rapid identification of cultivable oral treponemes. *Oral Microbiol Immunol*. 1999;14:117–21.
- Savitt ED, Strzempko MN, Vaccaro KK, Peros WJ, French CK. Comparison of cultural methods and DNA probe analyses for the detection of *Actinobacillus actinomycetemcomitans*, *Bacteroides gingivalis*, and *Bacteroides intermedius* in subgingival plaque samples. *J Periodontol*. 1988;59:431–8.
- Saygun I, Nizam N, Keskiner I, Bal V, Kubar A, Açikel C, Serdar M, Slots J. Salivary infectious agents and periodontal disease status. *J Periodontal Res*. 2011; 46:235–9.
- Scarfe WC. “All that glitters is not gold”: standards for cone-beam computerized tomographic imaging. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2011;111:402–8.
- Scarfe WC, Farman AG. What is cone-beam CT and how does it work? *Dent Clin North Am*. 2008;52:707–30.
- Scarfe WC, Farman AG, Sukovis P. Clinical applications of cone-beam computed tomography in dental practice. *J Can Dent Assoc*. 2006;72:75–80.
- Schara R, Sersa I, Skaleric U. T1 relaxation time and magnetic resonance imaging of inflamed gingival tissue. *Dentomaxillofac Radiol*. 2009;38:216–23.
- Schei O, Waerhaug J, Lovdal A, Arno A. Alveolar bone loss as related to oral hygiene and age. *J Periodontol*. 1959;30:7–16.
- Schütte UM, Abdo Z, Bent SJ, Shyu C, Williams CJ, Pierson JD, Forney LJ. Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. *Appl Microbiol Biotechnol*. 2008;80:365–80.
- Scott DA, Singer DL. Suppression of overt gingival inflammation in tobacco smokers – clinical and mechanistic considerations. *Int J Dent Hyg*. 2004;2: 104–10.
- Shaddox LM, Walker C. Microbial testing in periodontics: value, limitations and future directions. *Periodontol*. 2000. 2009;50:25–38.
- Shannon IL, Chauncey HH. A parotid fluid collection device with improved stability characteristics. *J Oral Ther Pharmacol*. 1967;4:93–7.
- Shapiro L, Goldman H, Bloom A. Sulcular exudate flow in gingival inflammation. *J Periodontol*. 1979;50:301–4.
- Sharma S, Rasool HI, Palanisamy V, Mathisen C, Schmidt M, Wong DT, Gimzewski JK. Structural-mechanical characterization of nanoparticle exosomes in human saliva, using correlative AFM, FESEM, and force spectroscopy. *ACS Nano*. 2010;4:1921–6.
- Shern RJ, Fox PC, Cain JL, Li S-H. A method for measuring the flow of saliva from the minor salivary glands. *J Dent Res*. 1990;69:1146–9.
- Shrout MK, Roberson B, Potter BJ, Mailhot JM, Hildebolt CF. A comparison of 2 patient populations using fractal analysis. *J Periodontol*. 1998;69:9–13.
- Silva-Boghossian CM, Amaral CS, Maia LC, Luiz RR, Colombo AP. Manual and electronic probing of the periodontal attachment level in untreated periodontitis: a systematic review. *J Dent*. 2008;36:651–7.
- Siqueira JF, Rôças IN, De Uzeda M, Colombo AP, Santos KR. Comparison of 16S rDNA-based PCR and checkerboard DNA-DNA hybridisation for detection of selected endodontic pathogens. *J Med Microbiol*. 2002;51:1090–6.
- Siqueira Jr JF, Rôças IN, Alves FR, Silva MG. Bacteria in the apical root canal of teeth with primary apical periodontitis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2009;107:721–6.
- Sixou M, Duffaut-Lagarigue D, Lodter JP. A comparison between 4 subgingival bacteriologic sampling technics (in French). *J Biol Buccale*. 1991;19:16–21.
- Slavkin HC, Fox CH, Meyer DM. Salivary diagnostics and its impact in dentistry, research, education, and the professional community. *Adv Dent Res*. 2011;23:381–6.
- Slots J, Hafström C, Rosling B, Dahlén G. Detection of *Actinobacillus actinomycetemcomitans* and *Bacteroides gingivalis* in subgingival smears by the indirect fluorescent antibody technique. *J Periodontal Res*. 1985;20:613–20.
- Smola SF, Rettenberger G, Simmet T, Burysek L. Comparison of sample collection methods for the PCR detection of oral anaerobic pathogens. *Lett Appl Microbiol*. 2003;36:101–5.
- Socransky SS, Smith C, Martin L, Paster BJ, Dewhirst FE, Levin AE. “Checkerboard” DNA-DNA hybridization. *Biotechniques*. 1994;17:788–92.
- Socransky SS, Haffajee AD, Smith C, Martin L, Haffajee JA, Uzel NG, Goodson JM. Use of checkerboard DNA-DNA hybridization to study complex microbial ecosystems. *Oral Microbiol Immunol*. 2004;19:352–62.
- Söder B, Airila Månsson S, Söder PO, Kari K, Meurman J. Levels of matrix metalloproteinases-8 and -9 with simultaneous presence of periodontal pathogens in gingival crevicular fluid as well as matrix metallopro-

- teinase-9 and cholesterol in blood. *J Periodontal Res.* 2006;41:411–7.
- Sorsa T, Mäntylä P, Rönkä H, Kallio P, Kallis GB, Lundqvist C, Kinane DF, Salo T, Golub LM, Teronen O, Tikanoja S. Scientific basis of a matrix metalloproteinase-8 specific chair-side test for monitoring periodontal and peri-implant health and disease. *Ann N Y Acad Sci.* 1999;878:130–40.
- Sorsa T, Tjäderhane L, Salo T. Matrix metalloproteinases (MMPs) in oral diseases. *Oral Dis.* 2004;10:311–8.
- Sorsa T, Tjäderhane L, Konttinen YT, Lauhio A, Salo T, Lee HM, Golub LM, Brown DL, Mäntylä P. Matrix metalloproteinases: contribution to pathogenesis, diagnosis and treatment of periodontal inflammation. *Ann Med.* 2006;38:306–21.
- Sorsa T, Hernández M, Leppilähti J, Munjal S, Netuschil L, Mäntylä P. Detection of gingival crevicular fluid MMP-8 levels with different laboratory and chair-side methods. *Oral Dis.* 2010;16:39–45.
- Sorsa T, Tervahartiala T, Leppilähti J, Hernandez M, Gamonal J, Tuomainen AM, Lauhio A, Pussinen PJ, Mäntylä P. Collagenase-2 (MMP-8) as a point-of-care biomarker in periodontitis and cardiovascular diseases. Therapeutic response to non-antimicrobial properties of tetracyclines. *Pharmacol Res.* 2011;63:108–13.
- Southard TE, Southard KA, Lee A. Alveolar process fractal dimension and postcranial bone density. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2001;91:486–91.
- Spielmann N, Wong DT. Saliva: diagnostics and therapeutic perspectives. *Oral Dis.* 2011;17:345–54.
- Squeri R, La Fauci V, Cannavò G, Lo Giudice G, Sindoni L. Identification of the microorganisms responsible for periodontopathy by Multiplex RT-PCR. *J Prev Med Hyg.* 2006;47:142–5.
- Srinath R, Acharya AB, Thakur SL. Salivary and gingival crevicular fluid melatonin in periodontal health and disease. *J Periodontol.* 2010;81:277–83.
- Stewart JE, Christenson PD, Maeder LA, Palmer MA. Reliability of filter-strip sampling of gingival crevicular fluid for volume determination using the Periotron. *J Periodontal Res.* 1993;28:227–30.
- Steyer R, Schmitt MJ. The effects of aggregation across and within occasions on consistency, specificity and reliability. *Methodika.* 1990;IV:58–94.
- Strand P, Palmer RM, Wilson RF. Sampling of subgingival plaque: a comparison of two methods using darkfield microscopy. *Oral Microbiol Immunol.* 1987;2:142–4.
- Su H, Gornitsky M, Velly AM, Yu H, Benarroch M, Schipper HM. Salivary DNA, lipid, and protein oxidation in non-smokers with periodontal disease. *Free Radic Biol Med.* 2009;46:914–21.
- Surna A, Kubilius R, Sakalauskiene J, Vitkauskienė A, Jonaitis J, Saferis V, Gleiznys A. Lysozyme and microbiota in relation to gingivitis and periodontitis. *Med Sci Monit.* 2009;15:CR66–73.
- Taba Jr M, Kinney J, Kim AS, Giannobile WV. Diagnostic biomarkers for oral and periodontal diseases. *Dent Clin North Am.* 2005;49:551–71.
- Takagi Y, Sumi M, Sumi T, Ichikawa Y, Nakamura T. MR microscopy of the parotid glands in patients with Sjögren's syndrome: quantitative MR diagnostic criteria. *AJNR Am J Neuroradiol.* 2005;26:1207–14.
- Takane M, Sugano N, Iwasaki H, Iwano Y, Shimizu N, Ito K. New biomarker evidence of oxidative DNA damage in whole saliva from clinically healthy and periodontally diseased individuals. *J Periodontol.* 2002;73:551–4.
- Takeshita T, Nakano Y, Kumagai T, Yasui M, Kamio N, Shibata Y, Shiota S, Yamashita Y. The ecological proportion of indigenous bacterial populations in saliva is correlated with oral health status. *ISME J.* 2009;3:65–78.
- Tanner AC, Goodson JM. Sampling of microorganisms associated with periodontal disease. *Oral Microbiol Immunol.* 1986;1:15–22.
- Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M. A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. *Methods.* 2010;50:S1–5.
- Teles FR, Haffajee AD, Socransky SS. The reproducibility of curet sampling of subgingival biofilms. *J Periodontol.* 2008;79:705–13.
- Teles RP, Likhari V, Socransky SS, Haffajee AD. Salivary cytokine levels in subjects with chronic periodontitis and in periodontally healthy individuals: a cross-sectional study. *J Periodontal Res.* 2009;44:411–7.
- Teles R, Sakellari D, Teles F, Konstantinidis A, Kent R, Socransky S. Haffajee Relationships among gingival crevicular fluid biomarkers, clinical parameters of periodontal disease, and the subgingival microbiota. *J Periodontol.* 2010a;81:89–98.
- Teles RP, Gursky LC, Faveri M, Rosa EA, Teles FR, Feres M, Socransky SS, Haffajee AD. Relationships between subgingival microbiota and GCF biomarkers in generalized aggressive periodontitis. *J Clin Periodontol.* 2010b;37:313–23.
- Tessier JF, Ellen RP, Birek P, Kulkarni GV, McCulloch CA. Relationship between periodontal probing velocity and gingival inflammation in human subjects. *J Clin Periodontol.* 1993;20:41–8.
- Thorat Manojkumar S, Pradeep AR, Garg G, Raju A. Gingival crevicular fluid levels of oncostatin M in periodontal conditions. *Cytokine.* 2010;50:248–52.
- Thurnheer T, Gmur R, Giertsen E, Guggenheim B. Automated fluorescent in situ hybridization for the specific detection and quantification of oral streptococci in dental plaque. *J Microbiol Methods.* 2001;44:39–47.
- Toker H, Poyraz O, Eren K. Effect of periodontal treatment on IL-1beta, IL-1ra, and IL-10 levels in gingival crevicular fluid in patients with aggressive periodontitis. *J Clin Periodontol.* 2008;35:507–13.
- Totan A, Greabu M, Totan C, Spinu T. Salivary aspartate aminotransferase, alanine aminotransferase and alkaline phosphatase: possible markers in periodontal diseases? *Clin Chem Labor Med.* 2006;44:612–5.
- Tözüm TF, Hatipoğlu H, Yamalik N, Gürsel M, Alptekin NO, Ataoğlu T, Marakoğlu I, Gürsoy UK, Eratalay K. Critical steps in electronic volume quantification of

- gingival crevicular fluid: the potential impact of evaporation, fluid retention, local conditions and repeated measurements. *J Periodontol Res.* 2004;39:344–57.
- Tran SD, Rudney JD. Multiplex PCR using conserved and species-specific 16SrRNA gene primers for simultaneous detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. *J Clin Microbiol.* 1996;34:2674–8.
- Tran SD, Rudney JD. Improved multiplex PCR using conserved and species-specific 16S rRNA gene primers for simultaneous detection of *Actinobacillus actinomycetemcomitans*, *Bacteroides forsythus*, and *Porphyromonas gingivalis*. *J Clin Microbiol.* 1999;37:3504–8.
- Trikilis N, Rawlinson A, Walsh TF. Periodontal probing depth and subgingival temperature in smokers and non-smokers. *J Clin Periodontol.* 1999;26:38–43.
- Trindade SC, Gomes-Filho IS, Meyer RJ, Vale VC, Pugliese L, Freire SM. Serum antibody levels against *Porphyromonas gingivalis* extract and its chromatographic fraction in chronic and aggressive periodontitis. *J Int Acad Periodontol.* 2008;10:50–8.
- Tsai CY, Wolff LF, Germaine G, Hodges J. A rapid DNA probe test compared to culture methods for identification of subgingival plaque bacteria. *J Clin Periodontol.* 2003;30:57–62.
- Tsalikis L, Malaka E, Pavlitou E, Konstantinidis A. Aspartate aminotransferase levels in gingival crevicular fluid before and after initial periodontal treatment. *J Int Acad Periodontol.* 2001;3:68–74.
- Tuuta-Veselicky L, Famili P, Ceravolo FJ, Zullo T. A clinical study of an electronic constant force periodontal probe. *J Periodontol.* 1994;65:616–22.
- Türkoglu O, Emingil G, Kütükçüler N, Atilla G. Gingival crevicular fluid levels of cathelicidin LL-37 and interleukin-18 in patients with chronic periodontitis. *J Periodontol.* 2009;80:969–76.
- Tüter G, Kurtis B, Serdar M. Evaluation of gingival crevicular fluid and serum levels of high-sensitivity C-reactive protein in chronic periodontitis patients with or without coronary artery disease. *J Periodontol.* 2007;78:2319–24.
- Tyndall DA, Rathore S. Cone-beam CT diagnostic applications: caries, periodontal bone assessment, and endodontic applications. *Dent Clin North Am.* 2008;52:825–41.
- Tyndall DA, Clifton TL, Webber RL, Ludlow JB, Horton RA. TACT imaging of primary caries. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 1997;84:214–25.
- Uitto VJ. Gingival crevice fluid – an introduction. *Periodontol 2000.* 2003;31:9–11.
- Uitto VJ, Overall CM, McCulloch C. Proteolytic host cell enzymes in gingival crevice fluid. *Periodontol 2000.* 2003;31:77–104.
- Umeda M, Contreras A, Chen C, Bakker I, Slots J. The utility of whole saliva to detect the oral presence of periodontopathic bacteria. *J Periodontol.* 1998;69:828–33.
- Udpiké SX, Nowzari H. Fractal analysis of dental radiographs to detect periodontitis-induced trabecular changes. *J Periodontol Res.* 2008;43:658–64.
- Urdea MS, Neuwald PD, Greenberg BL, Glick M, Galloway J, Williams D, Wong DT. Saliva, diagnostics, and dentistry. *Adv Dent Res.* 2011;23:353–9.
- Van der Velden U, de Vries IH. Introduction of a new periodontal probe: the pressure probe. *J Clin Periodontol.* 1978;5:188–97.
- Van der Zee E, Davies EH, Newman HN. Marking width, calibration from tip and tine diameter of periodontal probes. *J Clin Periodontol.* 1991;18:516–20.
- Van der Weijden GA, Timmerman MF, Reijerse E, Wolffe GN, Van Winkelhoff AJ, Van der Velden U. The prevalence of *A. actinomycetemcomitans*, *P. gingivalis* and *P. intermedia* in selected subjects with periodontitis. *J Clin Periodontol.* 1994;21:583–8.
- van Weringh M, Barendregt DS, Rosema NA, Timmerman MF, van der Weijden GA. A thin or thick probe handle: does it make a difference? *Int J Dent Hyg.* 2006;4:140–4.
- Vandenberghe B, Jacobs R, Yang J. Diagnostic validity (or acuity) of 2D CCD versus 3D CBCT-images for assessing periodontal breakdown. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2007;104:395–401.
- Wahlfors J, Meurman JH, Väisänen P, Alakuijala P, Korhonen A, Torkko H, Jänne J. Simultaneous detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis* by a rapid PCR method. *J Dent Res.* 1995;74:1796–801.
- Warnakulasuriya S, Dietrich T, Bornstein MM, Casals Peidró E, Preshaw PM, Walter C, Wennström JL, Bergström J. Oral health risks of tobacco use and effects of cessation. *Int Dent J.* 2010;60:7–30.
- Watts T. Constant force probing with and without a stent in untreated periodontal disease: the clinical reproducibility problem and possible sources of error. *J Clin Periodontol.* 1987;14:407–11.
- Watts TL. Probing site configuration in patients with untreated periodontitis. A study of horizontal positional error. *J Clin Periodontol.* 1989;16:529–33.
- Webber RL, Horton RA, Underhill TE, Ludlow JB, Tyndall DA. Comparison of film, direct digital, and tuned-aperture computed tomography images to identify the location of crestal defects around endosseous titanium implants. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 1996;81:480–90.
- Wei F, Patel P, Liao W, Chaudhry K, Zhang L, Arellano-Garcia M, Hu S, Elashoff D, Zhou H, Shukla S, Shah F, Ho CM, Wong DT. Electrochemical sensor for multiplex biomarkers detection. *Clin Cancer Res.* 2009;15:4446–52.
- Weinberg MA, Eskow RN. Periodontal terminology revisited. *J Periodontol.* 2003;74:563–5.
- Weinberg M, Westphal C, Froum S, Palat M, Schoor R. Comprehensive periodontics for the dental hygienist. 3rd ed. Boston: Pearson; 2010.
- Wilczynska-Borawska M, Borawski J, Kovalchuk O, Chyczewski L, Stokowska W. Hepatocyte growth factor in saliva is a potential marker of symptomatic periodontal disease. *J Oral Sci.* 2006;48:47–50.
- Winter AA. Measurement of the millimeter markings of periodontal probes. *J Periodontol.* 1979;50:483–5.

- Wolf DL, Lamster IB. Contemporary concepts in the diagnosis of periodontal disease. *Dent Clin North Am*. 2011;55:47–61.
- Wolff LF, Koller NJ, Smith QT, Mathur A, Aeppli D. Subgingival temperature: relation to gingival crevicular fluid enzymes, cytokines, and subgingival plaque micro-organisms. *J Clin Periodontol*. 1997;24:900–6.
- Xiang X, Sowa MG, Iacopino AM, Maev RG, Hewko MD, Man A, Liu KZ. An update on novel non-invasive approaches for periodontal diagnosis. *J Periodontol*. 2010a;81:186–98.
- Xiang XM, Liu KZ, Man A, Ghiabi E, Cholakias A, Scott DA. Periodontitis-specific molecular signatures in gingival crevicular fluid. *J Periodontal Res*. 2010b;45:345–52.
- Xu L, Yu Z, Lee HM, Wolff MS, Golub LM, Sorsa T, Kuula H. Characteristics of collagenase-2 from gingival crevicular fluid and peri-implant sulcular fluid in periodontitis and peri-implantitis patients: pilot study. *Acta Odontol Scand*. 2008;66:219–24.
- Yakob M, Kari K, Tervahartiala T, Sorsa T, Söder PÖ, Meurman JH, Söder B. Associations of periodontal microorganisms with salivary proteins and MMP-8 in gingival crevicular fluid. *J Clin Periodontol*. 2012;39:256–63.
- Yakob M, Kari K, Tervahartiala T, Sorsa T, Söder PO, Meurman JH, Söder B. Associations of periodontal microorganisms with salivary proteins and MMP-8 in gingival crevicular fluid. *J Clin Periodontol*. 2012;39:256–63.
- Yetkin Ay Z, Sütçü R, Uskun E, Bozkurt FY, Berker E. The impact of the IL-11: IL-17 ratio on the chronic periodontitis pathogenesis: a preliminary report. *Oral Dis*. 2009;15:93–9.
- Yin X, Bunn CL, Bartold PM. Detection of tissue plasminogen activator (t-PA) and plasminogen activator inhibitor 2(PAI-2) in gingival crevicular fluid from healthy, gingivitis and periodontitis patients. *J Clin Periodontol*. 2000;27:149–56.
- Yoo JY, Kim HC, Zhu W, Kim SM, Sabet M, Handfield M, Hillman J, Progulsk-Fox A, Lee SW. Identification of *Tannerella forsythia* antigens specifically expressed in patients with periodontal disease. *FEMS Microbiol Lett*. 2007;275:344–52.
- Yoshida A, Kawada M, Suzuki N, Nakano Y, Oho T, Saito T, Yamashita Y. TaqMan real-time polymerase chain reaction assay for the correlation of *Treponema denticola* numbers with the severity of periodontal disease. *Oral Microbiol Immunol*. 2004;19:196–200.
- Yoshie H, Tai H, Kobayashi T, Oda-Gou E, Nomura Y, Numabe Y, Ito K, Kurihara H, Kamoi K. Salivary enzyme levels after scaling and interleukin-1 genotypes in Japanese patients with chronic periodontitis. *J Periodontol*. 2007;78:498–503.
- Yoshihara A, Deguchi T, Hanada N, Miyazaki H. Relation of bone turnover markers to periodontal disease and jaw bone morphology in elderly Japanese subjects. *Oral Dis*. 2009;15:176–81.
- Yoshinari N, Kawase H, Mitani A, Ito M, Sugiishi S, Matsuoka M, Shirozu N, Ishihara Y, Bito B, Hiraga M, Arakawa K, Noguchi T. Effects of scaling and root planing on the amounts of interleukin-1 and interleukin-1 receptor antagonist and the mRNA expression of interleukin-1beta in gingival crevicular fluid and gingival tissues. *J Periodontal Res*. 2004;39:158–67.
- Yoshino T, Laine ML, Van Winkelhoff AJ, Dahlén G. Genotype variation and capsular serotypes of *Porphyromonas gingivalis* from chronic periodontitis and periodontal abscesses. *FEMS Microbiol Lett*. 2007a;270:75–81.
- Yoshino T, Laine ML, Van Winkelhoff AJ, Dahlén G. Genotypic characterization of *Porphyromonas gingivalis* isolated from Swedish patients with periodontitis and from periodontal abscesses. *Oral Microbiol Immunol*. 2007b;22:195–200.
- Zambon JJ, Reynolds HS, Chen P, Genco RJ. Rapid identification of periodontal pathogens in subgingival dental plaque. Comparison of indirect immunofluorescence microscopy with bacterial culture for detection of *Bacteroides gingivalis*. *J Periodontol*. 1985;56(Suppl):32–40.
- Zambon JJ. Principles of evaluation of the diagnostic value of subgingival bacteria. *Ann Periodontol*. 1997 Mar;2(1):138–48.
- Zhang L, Henson BS, Camargo PM, Wong DT. The clinical value of salivary biomarkers for periodontal disease. *Periodontol* 2000. 2009;51:25–37.
- Zheng P, Chen H, Shi S, Jepsen S, Eberhard J. Periodontal parameters and platelet-activating factor levels in serum and gingival crevicular fluid in a Chinese population. *J Clin Periodontol*. 2006;33:797–802.
- Zijnga V, van Leeuwen MB, Degener JE, Abbas F, Thurnheer T, Gmür R, Harmsen HJ. Oral biofilm architecture on natural teeth. *PLoS One*. 2010;5:e9321.
- Zybutz M, Rapoport D, Laurell L, Persson GR. Comparisons of clinical and radiographic measurements of inter-proximal vertical defects before and 1 year after surgical treatments. *J Clin Periodontol*. 2000;27:179–86.

A *periodontal diagnosis* is an important label that clinicians place on a patient's periodontal condition or disease. In the current practice of periodontics, it is primarily derived from information obtained from the patient's medical and dental histories combined with findings from a thorough oral examination. A carefully considered periodontal diagnosis is of major importance in the subsequent management of a patient's periodontal disease. An accurate diagnosis is often a first step towards development of a well-designed and appropriate treatment plan that when implemented leads to resolution of the patient's periodontal infection. An incorrect diagnosis often leads to an ill-conceived treatment approach that ultimately fails to resolve the patient's periodontal problem (Armitage 2004).

13.1 Types of Diagnostic Evaluation Studies

Before a new diagnostic test is introduced into medical practice, a series of studies may be undertaken to assess its performance and impact. Each study asks a separate question concerning the behaviour or impact of the test (Deeks 1999). Deeks (1999) proposed the following classification of diagnostic studies:

1. Studies of reliability

Reliability studies report on the consistency of test results when tests (or distinct parts of a testing procedure) are repeated under

a variety of conditions and circumstances. There are many sources of variability that can lead to poor reliability: observers may interpret images differently, testing equipment may vary between consecutive runs, and in some circumstances, natural measurement error may be too great to allow detection of real differences (Deeks 1999).

2. Studies of diagnostic accuracy or performance

Once it has been established that a test gives consistent values, the next stage is to investigate whether the test results actually relate to the disease in question. Studies of diagnostic accuracy or performance assess how well a test separates diseased from non-diseased patients (Deeks 1999).

Diagnostic accuracy studies have several unique features in terms of design which differ from standard intervention evaluations. Their aim is to determine how good a particular test is at detecting the target condition, and they usually have the following basic structure: A series of patients receive the test (or tests) of interest, known as the 'index test(s)' and also a reference standard. The results of the index test(s) are then compared to the results of the reference standard. The reference standard should be the best available method to determine whether or not the patient has the condition of interest. It may be a single test, clinical follow-up, or a combination of tests. Both the terms 'test' and 'condition' are interpreted in a broad sense. The term 'test' is used to refer to any procedure used to

gather information on the health status of an individual. This can include laboratory tests, surgical exploration, clinical examination, imaging tests, questionnaires, and pathology. Similarly, 'condition' can be used to define any health status including the presence of disease (e.g. influenza, alcoholism, depression, cancer), pregnancy, or different stages of disease (e.g. an exacerbation of multiple sclerosis) (Whiting et al. 2004).

Diagnostic accuracy studies allow the calculation of various statistics that provide an indication of 'test performance'—how good the index test is at detecting the target condition. These statistics include sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, diagnostics odds ratios, and receiver operating characteristic (ROC) curves (Whiting et al. 2004).

3. Studies of diagnostic and therapeutic process impact

Not all tests with good diagnostic properties are useful in practice. Before introduction into standard practice, a new test needs to demonstrate that it can beneficially alter the diagnostic process. The benefits could be that it produced more accurate diagnostic information, or the same information more quickly or cheaply, or that it obtains the same information in a manner which is less invasive, less risky, and more acceptable to patients. To study diagnostic impact, it is therefore necessary to compare aspects of the diagnoses made with and without the new test (Deeks 1999).

4. Studies of patient outcomes

Diagnostic tests can do harm as well as good. Wrong diagnoses are often harmful if they lead to delays in appropriate therapy, inappropriate therapy, or psychological morbidity. Treatments allocated on the basis of a correct diagnosis may have no effect, or could possibly do harm. In some instances, diagnostic tests themselves carry a morbidity or mortality risk. It is therefore possible that the potential negative impacts of using a test in clinical practice outweigh its benefits. This is especially important in screening programmes for

rare diseases where the vast majority of those receiving a test (those without the disease) can only be harmed by it (Deeks 1999).

13.2 Used Terminology

13.2.1 Diagnostic Versus Screening Tests

A clarification of the terminology of 'diagnostic test' and 'screening test' is first necessary, since the two are often confused and/or erroneously discussed interchangeably. At first glance, there may appear to be little difference between diagnostic tests and screening tests. In both cases, these will involve the application of clinical, laboratory, radiological, or other tests with the aim of determining whether a person is likely or unlikely to have a particular condition. Furthermore, similar measures of test performance can be calculated for both types of tests. Where they differ is in the purposes of the tests and the setting or context in which they are used (Needleman and Moles 2005).

13.2.1.1 Screening Test

The objective of a screening test is not to make a definitive diagnosis; rather, an attempt is made to categorize the person as being at high or low risk of a particular condition. Further diagnostic procedures are then required for those people who screen positive in order to determine their true status. Screening is carried out on people who are asymptomatic, some of whom will have the disease or risk factor of interest. There are two fundamental assumptions that underpin whether or not it is possible to screen for a condition. Firstly, that the condition can be detected before symptoms occur, that is, has a detectable preclinical phase (DPCP). Secondly, that the provision of an intervention at this DPCP stage is more beneficial than treating the person at a later stage. The intervention may be intended to be curative (if screening for a disease) or preventive (if screening for the presence of a risk factor) (Needleman and Moles 2005).

Table 13.1 Wilson and Jungner classic screening criteria (Andermann et al. 2008) (Reprinted with permission of the World Health Organization)

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a “once and for all” project.

Table 13.2 Synthesis of emerging screening criteria proposed over the past 40 years (Andermann et al. 2008) (Reprinted with permission of the World Health Organization)

- The screening programme should respond to a recognized need.
- The objectives of screening should be defined at the outset.
- There should be a defined target population.
- There should be scientific evidence of screening programme effectiveness.
- The programme should integrate education, testing, clinical services and programme management.
- There should be quality assurance, with mechanisms to minimize potential risks of screening.
- The programme should ensure informed choice, confidentiality and respect for autonomy.
- The programme should promote equity and access to screening for the entire target population.
- Programme evaluation should be planned from the outset.
- The overall benefits of screening should outweigh the harm.

Almost 40 years ago, the WHO commissioned a report on screening from James Maxwell Glover Wilson, then principal medical officer at the Ministry of Health in London, England, and Gunner Jungner, then chief of the Clinical Chemistry Department of Sahlgren’s Hospital in Gothenburg, Sweden. The report, published in 1968, was entitled: *Principles and practice of screening for disease*, and it has since become a public health classic. In their landmark publication, the authors were fundamentally preoccupied with the notion: ‘The central idea of early disease detection and treatment is essentially simple. However, the path to its successful achievement (on the one hand, bringing to treatment those with previously undetected disease, and, on the other, avoiding harm to those persons not in need of treatment) is far from simple though sometimes it may appear deceptively easy’ (Andermann et al. 2008).

***For this reason, Wilson and Jungner attempted to define screening criteria to guide the selec-

tion of conditions that would be suitable for screening, based, among other factors, on the capacity to detect the condition at an early stage and the availability of an acceptable treatment. They considered these criteria ‘especially important when case-finding is carried out by a public health agency, where the pitfalls may be more numerous than when screening is performed by a personal physician’ (Table 13.1) (Andermann et al. 2008).

Even when the same criteria are used, these have at times been criticized for being too vague or theoretical and thus difficult to assess in a consistent manner. Furthermore, even when criteria are met, there may still be logistical, social, or ethical reasons that preclude screening. Recently, several adaptations have been made to the classic criteria, and several new criteria have also emerged (Table 13.2). Many adaptations of classic criteria reflect issues raised by the growing interest in genetic screening (e.g. the importance of serious genetic conditions even if they are rare,

implications of genetic information for family members, need for analytical and clinical validity of screening tests, possibility of interventions that offer reproductive options, etc.). In contrast, many of the emerging criteria reflect broader trends that have shaped both Western medicine and society more generally over the past generation (e.g. increased consumerism, the shift away from paternalism towards informed choice, a focus on evidence-based health care, and the rise of managed care models that emphasize cost-effectiveness, quality assurance, and accountability of decision-makers) (Andermann et al. 2008).

13.2.1.2 Diagnostic Test

Here, the aim is to establish an actual diagnosis. Unlike screening, diagnostic tests are often motivated by the presence of signs and/or symptoms of disease (Needleman and Moles 2005).

13.2.2 Precision Versus Reliability and Accuracy Versus Validity

There are many articles in which the reliability of a scale is referred to as its ‘precision’ and validity as its ‘accuracy’. This is often illustrated with a diagram of a target pierced by some bullet holes; the tightness of the pattern of holes is a reflection of precision, and how close the centre of the pattern is to the target’s bullseye indicates the accuracy. For example, **Fig. 13.1a** uses this convention to show a test that is neither precise (there is a wide scatter) nor accurate (the holes are ‘biased’ towards the left). In **Fig. 13.1b**, there is still wide scatter, but the holes are spread relatively symmetrically around the centre (accurate, but with poor precision). There is more precision in **Fig. 13.1c**, but the test is inaccurate, while in **Fig. 13.1d**, the measure is both precise and accurate (Streiner and Norman 2006). Streiner and Norman (2006) showed that these terms are neither precise nor accurate when used in this way.

Precision is usually defined as the degree to which a score obtained by a person on one occasion is repeated on a second occasion (i.e. *test–retest reliability* in more traditional terms), or a score given by one rater is matched by that given by a second

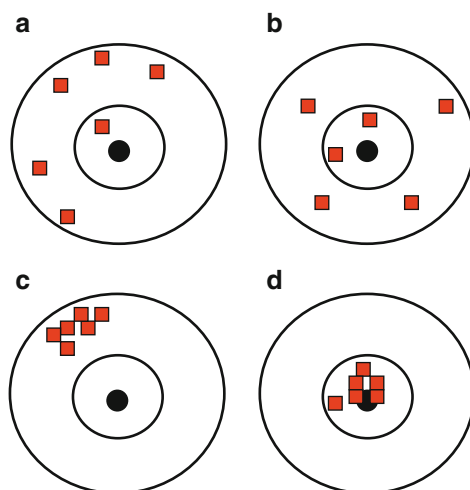


Fig. 13.1 Depiction of (a) low precision and low accuracy, (b) low precision and high accuracy, (c) high precision and low accuracy, and (d) high precision and accuracy (Streiner and Norman 2006; Reprinted with permission from Elsevier)

rater (i.e. *inter-rater reliability*). These definitions of reliability are fine as far as they go, but unfortunately, they do not go far enough. That is, reliability consists of more than just precision: scales can demonstrate high test–retest or inter-rater **reliability** (i.e. they are ‘precise’) but still be unreliable in certain circumstances, and ‘imprecise’ scales can still show good reliability. Further, ‘**accuracy**’ as a synonym for **validity** reflects an outdated conceptualization of validity, which has been superseded by one that emphasizes that validity tells us what conclusions can be drawn about a person based on a test result (Streiner and Norman 2006).

13.3 The Differences Between Studies of Diagnosis and Therapy

Diagnostic tests are different from therapeutic trials or evaluations. In studies of diagnostic accuracy, results from one or more tests are compared with the results obtained with the reference standard on the same subjects. Several factors threaten the internal and external validity of a study of diagnostic accuracy. These factors are all different from the quality of randomized or observational studies. Some of these factors have to do with the design of

such studies, others with the selection of patients, the execution of the tests, or the analysis of the data. Consequently, systems to grade the strength of scientific evidence assessed a multitude of articles and developed separate systems for randomized trials, observational studies, and diagnostic studies. Among differences between important quality features of RCTs and diagnostic studies should be noted that while randomization is the major criteria for a therapeutic trial, appropriate criterion or a gold standard is the major and most important criteria for a diagnostic study. It has also observed that observational studies are generally the most appropriate for answering questions related to diagnostic accuracy (Manchikanti et al. 2009a).

There are also significant differences between meta-analyses of therapeutic intervention studies, which have already been addressed in published handbooks, and meta-analyses of prognostic factors or diagnostic tests which are more recent and less standardized than the first ones. Meta-analyses of studies comparing interventions or treatments usually include randomized studies with two similar groups assessing the same intervention, in general compared with placebo or with a conventional treatment. Meta-analyses of studies on prognostic factors or diagnostic tests, in turn, face different challenges, such as different cutoff points for the positive or negative result of a test, or assessment of tests that were performed in prospective studies for the analysis of therapeutic interventions (Sousa and Ribeiro 2009).

13.4 Defining and Interpreting the Accuracy Diagnostic Test

13.4.1 The Accuracy of a Test

When a clinician asks ‘is this test accurate?’ the real question of interest is ‘does this test give the right answer?’ From a statistical standpoint, there is not a mathematical expression for accuracy; in evaluating the accuracy of a test, one must be familiar with the sensitivity and specificity of the test. A test is highly accurate if it is both highly sensitive *and* highly specific. From a practical standpoint, a test is accurate if it gives the ‘right’

Table 13.3 A 2×2 contingency table or decision matrix for diagnostic and prognostic tests

	Disease present	Disease absent
Test positive	A (true-positive)	C (false-positive)
Test negative	B (false-negative)	D (true-negative)

answer more often than not, which requires comparison with some gold standard for diagnosis. Most of the time, such a gold standard is not available and the clinician must decide just how accurate the test must be in a given situation (Nirula and Brasel 2006).

13.4.2 Test Sensitivity and Specificity

The standard procedure for validating and determining the utility of diagnostic and prognostic tests is a construction of a **2×2 contingency table or decision matrix** (Table 13.3). The contingency table can be used for tests intended to detect the presence of an existing disease state—a truly diagnostic test—as well as for tests that are putative predictors of continuing or future disease, prognostic tests. Evaluation of the diagnostic test is conducted in a cross-sectional study, and measurements are made of the prevalence of the disease in the study sample. Prognostic tests, on the other hand, are evaluated longitudinally and measurement is made of the incidence of new cases (or sites) (Mandel 1997).

Sensitivity is the proportion of people (or sites) with true (or active) disease identified as positive by the test: $A/A + B$ (Altman and Bland 1994a).

Specificity is the proportion of people (or sites) without (inactive) disease determined as such by the test: $D/C + D$ (Altman and Bland 1994a).

The sensitivity and specificity are proportions, so confidence intervals can be calculated for them using standard methods for proportions (Altman and Bland 1994a).

A perfect test would have both a sensitivity and specificity of 1.00—a situation consummately to be desired but not achievable in medical and dental practice. Pragmatically, a diagnostic test is considered useful for periodontal examination if it has a sensitivity and specificity of 0.70 or greater (Mandel 1997).

13.4.3 Positive and Negative Predictive Value

The whole point of a diagnostic test is to use it to make a diagnosis, so we need to know the probability that the test will give the correct diagnosis. The sensitivity and specificity do not give us this information. Instead, we must approach the data from the direction of the test results, using predictive values (Altman and Bland 1994b).

Positive predictive value (PPV) is the proportion of patients with positive test results who are correctly diagnosed (Altman and Bland 1994b). Positive predictive value = $A/A + C$.

Negative predictive value (NPV) is the proportion of patients with negative test results who are correctly diagnosed (Altman and Bland 1994b). Negative predictive value = $D/B + D$.

The positive and negative predictive values (PPV and NPV) can be calculated for any prevalence as follows:

$$PPV = \text{sensitivity} \times \text{prevalence} / \text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})$$

$$NPV = \text{specificity} \times (1 - \text{prevalence}) / (1 - \text{sensitivity}) \times \text{prevalence} + \text{specificity} \times (1 - \text{prevalence})$$

If the prevalence of the disease is very low, the positive predictive value will not be close to 1 even if both the sensitivity and specificity are high. Thus, in screening the general population, it is inevitable that many people with positive test results will be false positives (Altman and Bland 1994b).

The prevalence can be interpreted as the probability before the test is carried out that the subject has the disease, known as the prior probability of disease. The positive and negative predictive values are the revised estimates of the same probability for those subjects who are positive and negative on the test and are known as posterior probabilities. The difference between the prior and posterior probabilities is one way of assessing the usefulness of the test (Altman and Bland 1994b).

An odds ratio can also be calculated from a contingency table: Odds ratio = AD/BC . It is a measure of the level of increased risk for developing or activating disease when the particular test is

positive. Odds ratio is most appropriately calculated in retrospective case-control studies or cross-sectional studies. In prospective studies, the characterization of the strength of the association between a particular risk factor, host or environmental, is described as the relative risk. Statistically, **the relative risk** is the ratio of the risk of developing disease in exposed individuals to the risk in the unexposed group (Mandel 1997).

13.4.4 Likelihood Ratios

In clinical practice, it is essential to know how a particular test result predicts the risk of abnormality. Sensitivities and specificities (Altman and Bland 1994a) do not do this: they describe how abnormality (or normality) predicts particular test results. Predictive values (Altman and Bland 1994b) do give probabilities of abnormality for particular test results, but depend on the prevalence of abnormality in the study sample and can rarely be generalized beyond the study (except when the study is based on a suitable random sample, as is sometimes the case for population screening studies). Likelihood ratios provide a solution as they can be used to calculate the probability of abnormality, while adapting for varying prior probabilities of the chance of abnormality from different contexts (Deeks and Altman 2004).

For any test result, we can compare the probability of getting that result if the patient truly had the condition of interest with the corresponding probability if he or she were healthy. The ratio of these probabilities is called the likelihood ratio, calculated as $\text{sensitivity}/(1 - \text{specificity})$ (Altman and Bland 1994b).

A likelihood ratio greater than 1 indicates that the test result is associated with the presence of the disease, whereas a likelihood ratio less than 1 indicates that the test result is associated with the absence of disease. The further likelihood ratios are from 1 the stronger the evidence for the presence or absence of disease. Likelihood ratios above 10 and below 0.1 are considered to provide strong evidence to rule in or rule out diagnoses,

respectively, in most circumstances. When tests report results as being either positive or negative, the two likelihood ratios are called the positive likelihood ratio and the negative likelihood ratio (Deeks and Altman 2004).

Likelihood ratios do not have the drawbacks of the aforementioned other measures of test performance. They are not affected by changes in the prevalence of disease and can be used when test results are grouped into more than two categories. Another useful property of likelihood ratios is that they can be converted into the ‘post-test probability’ of disease given knowledge of the ‘pretest probability’ of disease (Needleman and Moles 2005).

13.4.5 Post-test Probability of Disease

The final step in determining whether a patient has a particular disease given a positive or negative test result is to combine the likelihood ratio with the pretest probability of disease, or the prevalence of disease in a particular population. For a very rare or very common disease, the test must be very definitive to change the pretest probability of disease; the likelihood ratio must be very large or very small. For most diseases, positive likelihood ratios of near 10 and negative likelihood ratios near 0.1 will significantly change pretest probabilities. A nomogram using the pretest probability and the likelihood ratio to identify the post-test probability is shown in Fig. 13.2 (Nirula and Brasel 2006).

Mathematically, the post-test probability of disease can be expressed as:

Post-test probability of disease = pretest probability of disease (disease prevalence) × likelihood ratio (Nirula & Brasel, 2006).

13.4.6 Receiver Operating Curves (ROC)

We have previously considered diagnosis based on tests that give a yes or no answer (binary

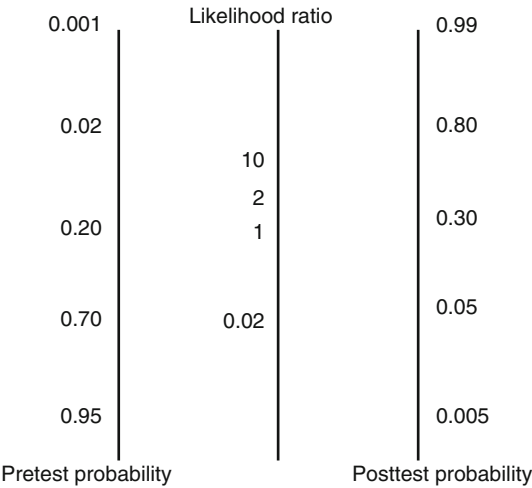


Fig. 13.2 Predicting post-test probability. Nomogram illustrates how to derive the post-test probability when the pretest probability and likelihood ratio are known. Most likelihood ratios are between 0.1 and 10; the closer a likelihood ratio is to 1, the less effect it has on changing the pretest probability. Note that probabilities must lie between 0 and 1; any probability that “multiplies” to a number higher than 1 is treated as a probability of 1 (Nirula and Brasel 2006; Reprinted with permission from Elsevier)

response). Many diagnostic tests, however, are quantitative, notably in clinical chemistry. The same statistical approach can be used only if we can select a cutoff point to distinguish ‘normal’ from ‘abnormal’, which is not a trivial problem. Firstly, we can investigate to what extent the test results differ among people who do or do not have the diagnosis of interest. The receiver operating characteristic (ROC) plot is one way to do this (Altman and Bland 1994c).

A receiver operating characteristic plot is obtained by calculating the sensitivity and specificity of every observed data value and plotting sensitivity against 1 – specificity, as in Fig. 13.3. A test that perfectly discriminates between the two groups would yield a ‘curve’ that coincided with the left and top sides of the plot. A test that is completely useless would give a straight line from the bottom left corner to the top right corner. In practice, there is virtually always some overlap of the values in the two groups, so the curve will lie somewhere between these extremes (Altman and Bland 1994c).

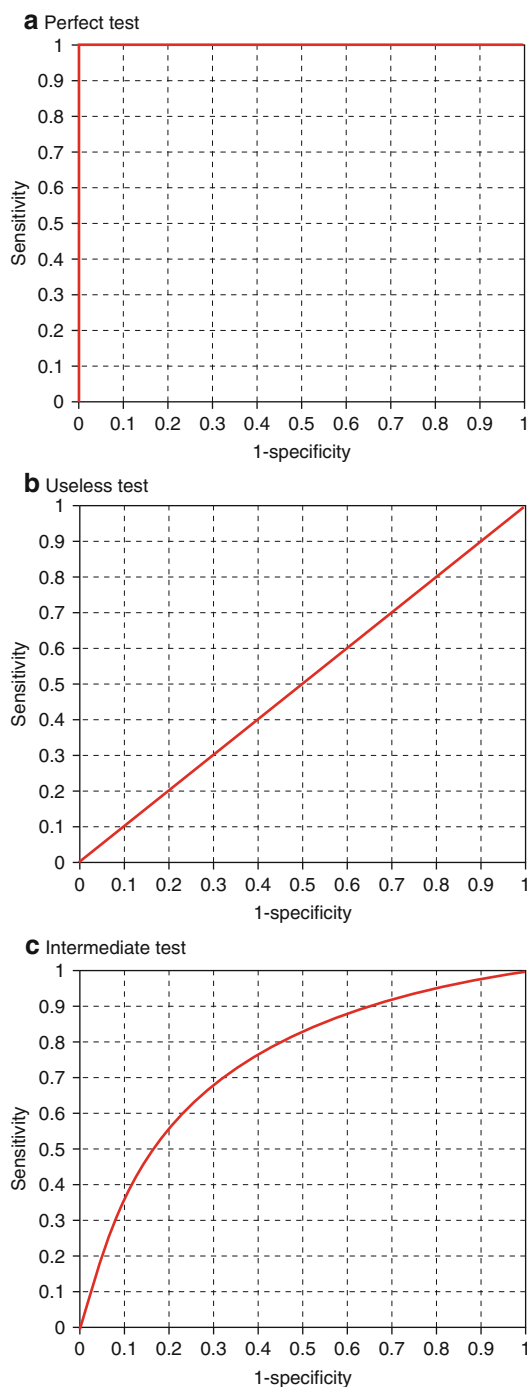


Fig. 13.3 Hypothetical ROC plots (Needleman and Moles 2005; Reprinted with permission from John Wiley & Sons, Inc.)

13.5 Guide for Reporting Studies of Diagnostic Tests: The STARD Initiative

Complete and accurate reporting of studies of diagnostic accuracy is necessary to enable readers to assess the potential for bias in the study and to evaluate the generalizability of the results. A group of statisticians and editors has developed the STARD (*Standards for Reporting of Diagnostic Accuracy*) statement to improve the reporting the quality of reporting of studies of diagnostic accuracy (Bossuyt et al. 2003a, b).

The statement consists of a checklist of 25 items (Table 13.4) and flow diagram (Fig. 13.4) that authors can use to ensure that all relevant information is present.

The items in the checklist and flow chart can help authors to describe essential elements of the design and conduct of the study, the execution of tests, and the results. The items are arranged under the usual headings of a medical research article, but this is not intended to dictate the order in which they have to appear within an article (Bossuyt et al. 2003c).

A flow diagram has the potential to communicate vital information about the design of a study and the flow of participants in a transparent manner. A comparable flow diagram has become an essential element in the CONSORT standards for reporting of randomized trials. The flow diagram could be even more essential in diagnostic studies, given the variety of designs employed in diagnostic research. Flow diagrams in the reports of studies of diagnostic accuracy indicate the process of sampling and selecting participants (external validity); the flow of participants in relation to the timing and outcomes of tests; the number of subjects who fail to receive the index test or the reference standard, or both (potential for verification bias); and the number of patients at each stage of the study, which provides the correct denominator for proportions (internal consistency) (Bossuyt et al. 2003c). Examples that reflect other designs appear on the STARD website (www.consortstatement.org/stardstatement.htm).

Table 13.4 STARD checklist for reporting diagnostic accuracy studies (Bossuyt et al. 2003c) (Reprinted with permission from BMJ Publishing Group Ltd.)

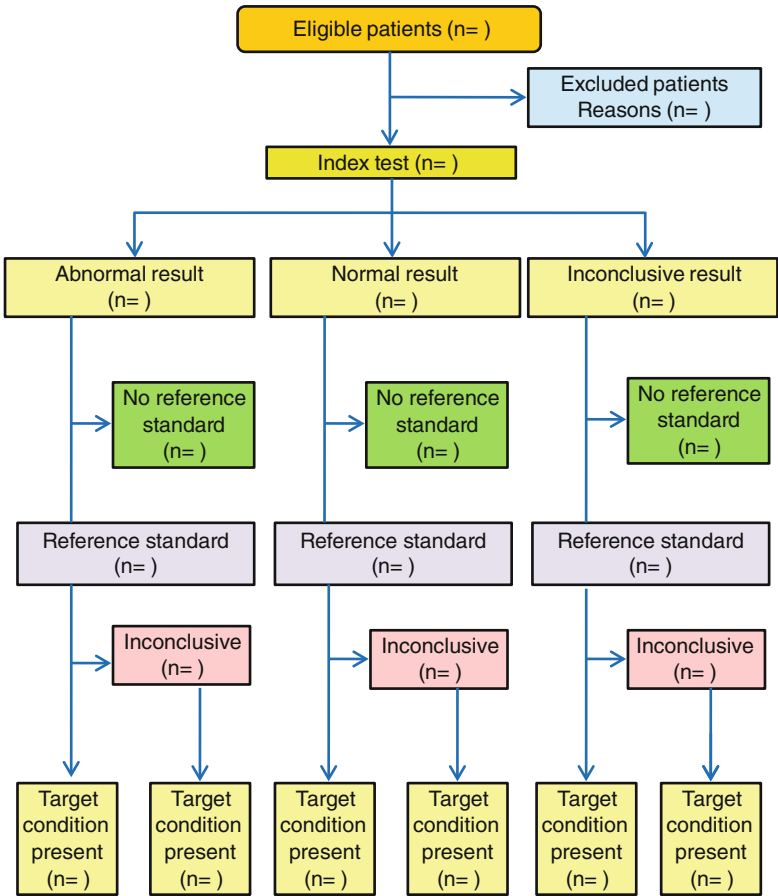
Section and topic	Item	Description
Title, abstract, and keywords	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity')
Introduction	2	State the research questions or aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups
Methods: Participants	3	Describe the study population: the inclusion and exclusion criteria and the settings and locations where the data were collected
	4	Describe participant recruitment: was this based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?
	5	Describe participant sampling: was this a consecutive series of participants defined by selection criteria in items 3 and 4? If not, specify how participants were further selected
	6	Describe data collection: was data collection planned before the index tests and reference standard were performed (prospective study) or after (retrospective study)?
Test methods	7	Describe the reference standard and its rationale
	8	Describe technical specifications of material and methods involved, including how and when measurements were taken, or cite references for index tests or reference standard, or both
	9	Describe definition of and rationale for the units, cut-off points, or categories of the results of the index tests and the reference standard
	10	Describe the number, training, and expertise of the persons executing and reading the index tests and the reference standard
	11	Were the readers of the index tests and the reference standard blind (masked) to the results of the other test? Describe any other clinical information available to the readers.
Statistical methods	12	Describe methods for calculating or comparing measures of diagnostic accuracy and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals)
	13	Describe methods for calculating test reproducibility, if done
Results: Participants	14	Report when study was done, including beginning and ending dates of recruitment
	15	Report clinical and demographic characteristics (eg age, sex, spectrum of presenting symptoms, comorbidity, current treatments, and recruitment centre)
	16	Report how many participants satisfying the criteria for inclusion did or did not undergo the index tests or the reference standard, or both; describe why participants failed to receive either test (a flow diagram is strongly recommended)
Test results	17	Report time interval from index tests to reference standard, and any treatment administered between
	18	Report distribution of severity of disease (define criteria) in those with the target condition and other diagnoses in participants without the target condition
	19	Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, report the distribution of the test results by the results of the reference standard
	20	Report any adverse events from performing the index test or the reference standard

(continued)

Table 13.4 (continued)

Section and topic	Item	Description
Estimates	21	Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals)
	22	Report how indeterminate results, missing responses, and outliers of index tests were handled
	23	Report estimates of variability of diagnostic accuracy between readers, centres, or subgroups of participants, if done
	24	Report estimates of test reproducibility, if done
Discussion	25	Discuss the clinical applicability of the study findings

Fig. 13.4 Prototype of a flow diagram for a study on diagnostic accuracy (Bossuyt et al. 2003c; Reprinted with permission from BMJ Publishing Group Ltd.)



13.6 Checklists for Assessing Studies on Diagnostic Test Accuracy

13.6.1 Bias and Variation in Studies of Diagnostic Test Accuracy

The validity will determine whether the results of a publication will be useful to patient management. The threats to validity have been divided into two

main categories: bias and variation (Lijmer et al. 1999; Whiting et al. 2004). **Bias** is a systematic deviation from the true result causing the estimate of test accuracy to be incorrect and may be a factor of design and/or conduct of a study. Often, this is in the direction of exaggerating the accuracy of a test. **Variation** in study design may lead to limitations in the generalizability of test findings. The main sources of bias and variation in the new study were demographic features, disease severity, disease

prevalence, partial verification bias, review bias, and observer variability (Needleman and Moles 2005; Whiting et al. 2004).

The impact of bias can be measured by comparing the diagnostic test accuracy in studies with ideal methodology and studies with flawed methodology (Needleman and Moles 2005). As a summary estimate, **the diagnostic odds ratio (DOR)**, the odds for a positive test result in diseased persons relative to the odds of a positive result in non-diseased persons, is calculated from a 2×2 table, incorporating sensitivity as well as specificity. Expressed in terms of sensitivity and specificity, the formula is:

$$\text{DOR} = \text{Sensitivity} / (1 - \text{Sensitivity}) : (1 - \text{Sensitivity}) / \text{Specificity}$$

The effect of study characteristics can be examined with a regression model that is adapted from the summary receiver operating characteristic curve model, developed for meta-analyses of diagnostic tests. Covariates can be added to this model to examine whether, on average, studies that failed to meet the methodological criteria yielded different DORs. The resulting parameter estimates of the covariates can be interpreted after antilogarithm transformation as **relative DORs (RDORs)**. They indicate the diagnostic performance of a test in studies failing to satisfy the methodological criterion, relative to its performance in studies with the corresponding feature. If the RDOR is larger than 1, studies not satisfying the criterion yield larger estimates of the DOR than studies with this corresponding feature (Lijmer et al. 1999).

13.6.2 Checklists for Quality Assessment of Diagnostic Accuracy Studies

13.6.2.1 Critical Appraisal Skills Programme (CASP) for Diagnostic Test

Critical Appraisal Skills Programme (CASP) has helped to develop an evidence-based approach in health and social care and aims to enable individuals to develop the skills to find and make

sense of research evidence, helping them to put knowledge into practice.

The CASP appraisal tool for diagnostic test studies is available free at www.casp-uk.net. According to CASP, three broad issues are considered when appraising a diagnostic test:

1. Are the results of the study valid?
2. What are the results?
3. Will the results help me and my patients/population?

The 12 questions are designed to help you think through these issues systematically and were adapted from Jaeschke et al. (1994). The first two questions are screening questions and can be answered quickly:

- Was there a clear question for the study to address?
- Was there a comparison with an appropriate reference standard?

If the answer to both is 'yes', it is worth proceeding with the remaining questions. You are asked to record a 'yes', 'no', or 'can't tell' to most of the questions. A number of italicized prompts are given after each question. These are designed to remind you why the question is important (Table 13.5).

13.6.2.2 QUADAS-2: Tool for the Quality Assessment of Diagnostic Accuracy Studies

Quality assessment is as important in systematic reviews of diagnostic accuracy studies as it is in any other review. Since its publication in 2003, the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool has been widely used (Whiting et al. 2003). More than 200 review abstracts in the Database of Abstracts of Reviews of Effects mention this tool, and it has been cited more than 500 times. The QUADAS tool is recommended for use in systematic reviews of diagnostic accuracy by the Agency for Healthcare Research and Quality, Cochrane Collaboration, and the UK National Institute for Health and Clinical Excellence (Whiting et al. 2011).

The original QUADS tool was structured as a list of 14 questions which should each be answered 'yes', 'no', or 'unclear' (Whiting et al. 2003). Experience, anecdotal reports, and feedback suggested areas for improvement; therefore, QUADAS-2 was developed. The full QUADAS-2

Table 13.5 Critical Appraisal Skills Programme (CASP) for diagnostic test scale (Reprinted with permission, available at www.casp-uk.net)

(A) Are the results of the study valid?	1. Was there a clear question for the study to address?	Yes ☐	Can't tell ☐	No ☐
	2. Was there a comparison with an appropriate reference standard?	Yes ☐	Can't tell ☐	No ☐
	3. Did all patients get the diagnostic test and the reference standard?	Yes ☐	Can't tell ☐	No ☐
	4. Could the results of the test of interest have been influenced by the results of the reference standard?	Yes ☐	Can't tell ☐	No ☐
	5. Is the disease status of the tested population clearly described?	Yes ☐	Can't tell ☐	No ☐
	6. Were the methods for performing the test described in sufficient detail?	Yes ☐	Can't tell ☐	No ☐
(B) If so, what are the results?	7. What are the results?			
	8. How sure are we about these results?			
(C) Will the results help me and my patients/population?	9. Can the results be applied to your patients/the population of interest?	Yes ☐	Can't tell ☐	No ☐
	10. Can the test be applied to your patient or population of interest?	Yes ☐	Can't tell ☐	No ☐
	11. Were all outcomes important to the individual or population considered?	Yes ☐	No ☐	
	12. What would be the impact of using this test on your patients/population?			

tool is available from the QUADAS website (www.quadas.org) (Supplement, available at www.annals.org). This tool is designed to assess the quality of primary diagnostic accuracy studies; it is not designed to replace the data extraction process of the review and should be applied in addition to extracting primary data (e.g. study design and results) for use in the review. The QUADAS tool consists of four key domains that discuss patient selection, index test, reference standard, and flow of patients through the study and timing of the index tests and reference standard (flow and timing) (**Table 13.1**) (Whiting et al. 2011).

The tool is completed in four phases: report the review question, develop review-specific guidance, review the published flow diagram for the primary study or construct a flow diagram if none is reported, and judge bias and applicability. Each domain is assessed in terms of the risk of bias, and the first three domains are also assessed in terms of concerns about applicability. Signalling questions are included to help judge the risk of bias; these questions flag aspects of

study design related to the potential for bias and aim to help reviewers judge risk of bias. This tool will allow for more transparent rating of bias and applicability of primary diagnostic accuracy studies (Whiting et al. 2011).

13.7 Systematic Reviews and Meta-analysis of Diagnostic Studies

Systematic reviews with meta-analysis of studies on diagnostic tests or prognostic factors are research tools that are still being developed. Since 1994, when a guideline for meta-analysis of studies on diagnostic tests (Irwig et al. 1994) was published, several different publications with criticisms and propositions on specific aspects of each stage of the process came up (Deville et al. 2002; Halligan 2005; Sousa and Ribeiro 2009). The use of meta-analysis for diagnostic and prognostic tests is still being developed, but it has become increasingly more important (Sousa and Ribeiro 2009).

13.7.1 How to Conduct Systematic Reviews and Meta-analyses of Diagnostic Accuracy Studies

Diagnostic test accuracy reviews face two major challenges. Firstly, they are limited by the quality and availability of primary test accuracy studies that address important relevant questions. More studies are needed which recruit suitable spectrums of participants, make direct comparisons between tests, use rigorous methodology, and clearly report their methods and findings. Secondly, more development is needed in the area of interpretation and presentation of the results of diagnostic test accuracy reviews (Leeflang et al. 2008).

Guidelines illustrating the evaluation of systematic reviews of diagnostic accuracy (Sousa and Ribeiro 2009; Leeflang et al. 2008) and meta-analysis (Irwig et al. 1994) delineate multiple steps in conducting a systematic review or meta-analysis.

13.7.1.1 The Cochrane Collaboration (Leeflang et al. 2008)

Leeflang et al. (2008) reviewed methodologic developments concerning problem formulation, location of literature, quality assessment, and meta-analysis of diagnostic accuracy studies by using their experience from the work on the Cochrane Handbook (Manchikanti et al. 2009b). Leeflang et al. (2008) described multiple objectives of the methodology for conducting systematic reviews of diagnostic test accuracy studies.

1. Definition of the objectives of the review

Diagnostic test accuracy refers to the ability of that test to distinguish between patients with disease (or more generally, a specified target condition) and those without. In such a test accuracy study, the results of the test under evaluation, or 'index test', are compared with those of the reference standard determined in the same patients. The reference standard is the best available method for identifying patients that have the target condition (Leeflang et al. 2008).

Test accuracy is not a fixed property of a test. It can vary between patient subgroups, with their spectrum of disease, with the clinical setting, and with the test interpreters and may depend on the results of prior testing. For this

reason, it is essential to include these elements in the study question (Leeflang et al. 2008).

2. Identification and selection of studies

A broad bibliographic search with no preset strategy is recommended. Restrictive electronic search filters are discouraged (Leeflang et al. 2008).

3. Assessment of methodological quality

More variability among diagnostic accuracy study results is to be expected than with randomized trials. Some of this variability is due to chance, as many diagnostic studies have small sample sizes. The remaining heterogeneity may be due to differences in study populations, but differences in study methods are also likely to result in differences in accuracy estimates. Test accuracy studies with design deficiencies can produce biased results (Leeflang et al. 2008).

Quality assessment of individual studies in systematic reviews is therefore necessary to identify potential sources of bias and to limit the effects of these biases on the estimates and the conclusions of the review. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist to assess the quality of diagnostic test accuracy studies was recommended (Leeflang et al. 2008).

4. Analysing the data and presenting the results

Statistical analyses using appropriate techniques, including summary ROC curves that take into account heterogeneity when meta-analysis is possible, should be used (Leeflang et al. 2008).

5. Interpretation of the results

The interpretation of the results offered in the systematic review should help readers to understand the implications for practice. This interpretation should consider whether evidence derived from the review suitably addresses the objectives of the review. It may involve considerations about whether the study sample was representative, whether the included studies indeed investigated the intended future role of the test under evaluation, and whether the results are unlikely to be biased. The potential effects of quality differences on the results, or the lack of high-quality studies should be considered. The interpretation of the findings should furthermore consider the consequences

of the false-positive and false-negative test results and whether the estimates of accuracy that were found are sufficiently high for the foreseen role that the test will have in practice. Some reviews may not result in useful summary estimates of sensitivity and specificity, for example, because of large variability in the individual study estimates, or because the authors only investigated the comparative accuracy by comparing summary ROC curves. A decision model could be used to structure the interpretation of the findings. Such a model would incorporate important factors as the disease prevalence, likely outcomes, and the available diagnostic and therapeutic interventions that may follow the test. Additional information, such as costs or important trade-offs between harms and benefits can be included (Leeftang et al. 2008).

All these details of methodology are described in detail in the Cochrane Handbook for Diagnostic Test Accuracy Reviews (<http://srdta.cochrane.org/handbook-dta-reviews>).

13.7.1.2 Tutorial for Systematic Review and Meta-analysis of Diagnostic Tests Studies

Sousa and Ribeiro (2009) described the methodology of systematic review and meta-analysis of the diagnostic tests studies, step by step. A literature review on the subject was made, the recommendations were compiled, and the paper was organized in:

- (a) Introduction
- (b) Details on the eight steps to be followed:
 1. Define clearly the question to be formulated.
 2. Search all reliable studies addressing the question in different sources.
 3. Select the studies by means of clear inclusion and exclusion criteria and evaluate the quality of these studies.
 4. Extract data from each study and display them clearly.
 5. Evaluate heterogeneity among the studies. The methodological aspects to be evaluated are shown in **Table 13.6**.

6. Calculate the results of each study (and combine them, if appropriate), estimating diagnostic accuracy (**Table 13.7**).
7. Assess the effect of variation in study validity on the estimates of diagnostic accuracy.
8. Interpret the results, assessing how much of the review and/or meta-analysis can be generalized according to the patients' characteristics.
- (c) Form of publication of a systematic review with meta-analysis.
- (d) Conclusion

The systematic review methods were thoroughly described with a critical analysis of the methods of statistical compilation of results, with emphasis on the utilization of the summary receiver operator characteristic curve. References for the details of each statistical technique used were provided in the meta-analysis. The authors concluded that systematic reviews with meta-analysis of diagnostic tests are useful in data compilation of various studies on the same subject, since they reduce biases and increase the statistical power of the primary research (Sousa and Ribeiro 2009).

13.7.1.3 Guidelines for Meta-analyses Evaluating Diagnostic Tests

To help researchers do and readers assess meta-analyses of diagnostic accuracy, Irwig et al. (1994) suggested guidelines on how they should be conducted, reported, and critically appraised. The steps include:

1. Determination of the objective and scope of the meta-analysis
2. Retrieval of the relevant literature
3. Extraction and displacement of the data
4. Estimation of diagnostic accuracy
5. Assessment of the effect of variation in study validity on estimates of diagnostic accuracy
6. Assessment of the effect of variation in the characteristics of patients and test on estimates of diagnostic accuracy (generalizability).

Their guidelines are based on current concepts of how to assess diagnostic tests and conduct meta-analyses. Attention to these guidelines can greatly improve our ability to synthesize

Table 13.6 List of aspects to be checked in the assessment of diagnostic and prognostic studies during the systematic review and meta-analysis (Sousa and Ribeiro 2009) and references therein (Reprinted under the terms of the Creative Commons Attribution License)

Age and gender distribution of the population studied.
Inclusion date and follow-up period of the study
Standardized reference test, adequacy of the gold standard chosen, evaluating whether this does not lead to the wrong classification of disease status.
Technical aspects of the performance of the test.
Evaluate the degree of missing data.
Original false and true-positive results, false and true-negative results. Occasionally, these data can be estimated from the sensitivity and specificity values as well as from the positive and negative values of the endpoint or reference test.
Reference values for the gold-standard test and for the index test, in a clear way and representative of the disease of interest.
The confidence interval and the standard error for test accuracy measurements.
The number of readers and their training for the index and the gold-standard test.
Presence of review bias: verify whether the test result in the study was evaluated blind to the endpoints and other tests (independent interpretation).
Presence of verification bias: the reference test may have been performed preferably in patients with positive tests, which is more frequent when the tests considered as a gold standard are invasive. In this case, the choice of patients for verification by the gold-standard test is not random.
Whether the reference test was performed in all patients. If the index and the gold-standard tests have not been performed in all patients, which is ideal, evaluate whether the choice of patients for the tests was random, thus decreasing the chance of bias.
Presence of clinical spectrum bias: lack of representation of the clinical spectrum of the disease of interest in the study population. Evaluate patients' demographic and clinical data such as age, gender, race, clinical characteristics, presence of symptoms, disease stage, duration, and comorbidities. The prevalence of the condition among the population studied provides a broader view of the spectrum, circumstances and potential of generalizability.
In screening tests, there may be excess diagnosis bias (when a disease that could progress asymptotically is detected), excess representation bias (for diseases that progress slowly, making them 'stand out' because of the screening), and early detection bias (which overestimates the effects of clinical benefits).

Table 13.7 Forms of summarizing test accuracy by means of meta-analysis (Sousa and Ribeiro 2009) (Reprinted under the terms of the Creative Commons Attribution License)

1. Combination of sensitivities and specificities
2. Combination of positive and negative likelihood ratios
3. Combination of diagnostic odds ratios
4. Diagnostic effectiveness scores (or effect size measure)
5. sROC curves (summary ROC or ordinary ROC curve)

information in the growing literature on diagnostic test evaluation. Meta-analysis can potentially provide a better understanding by examining the variability in estimates of diagnostic accuracy from primary studies, exploring whether this variability is explained by differences in study valid-

ity, and determining whether diagnostic accuracy based on the most valid studies varies for different clinical subgroups or categories of patients. By improving the conduct and reporting of primary studies, meta-analysis will also become a more valuable technique (Irwig et al. 1994).

13.7.1.4 Reporting of Systematic Reviews of Diagnostic Accuracy Studies

Even though Quality of Reporting of Meta-analyses QUOROM and Meta-analysis of Observational Studies in Epidemiology (MOOSE) have described proposals for reporting systematic reviews and meta-analysis, no such guidelines have been established for systematic reviews of

diagnostic accuracy studies. Further, diagnostic accuracy studies may also be reported in conjunction with evaluation of treatment. Consequently, based on QUOROM and MOOSE, a checklist was proposed for diagnostic accuracy studies by Manchikanti et al. (2009b) (Table 13.8).

Sousa and Ribeiro (2009) suggested by analogy with the Quality of Reporting of Meta-analysis (QUOROM) conference for publication of meta-analyses of studies on therapeutic intervention that the methodology should be thoroughly described when publishing the results of meta-analysis of diagnostic studies and each phase of the process should be explicit:

- The title should identify the study as a meta-analysis or systematic review.
- The summary should be structured with description of the following aspects: the clinical question, the sources and database, the methods of review and selection of the literature and of quantitative synthesis of the data in a reproducible form, the results with estimates and confidence intervals, and the conclusion with the main results.
- The introduction should contextualize and provide the background to the objective.
- The methodology should give details on the sources and search strategies, the period and language, criteria of study selection, form of assessment of publication bias, assessment of quality and methodological validity of the studies, the form of data extraction ideally by two researchers, the study characteristics, the form of assessment of heterogeneity, and the form of mathematically summarizing data.
- The results should present the review flow, the study characteristics assessing age and gender distribution, form of diagnosis of patient selection, relevant covariates, follow-up period, sample size, and the estimates of diagnostic or prognostic accuracy with the respective confidence intervals.
- In the discussion, summarize the key issues, discuss the clinical inferences based on the internal and external validity, interpret the results in light of all the evidences, describe the limitations and potential biases—especially

publication bias—and suggest further studies (Sousa and Ribeiro 2009).

13.8 Application of GRADE: Making Evidence-Based Recommendations About Diagnostic Tests in Clinical Practice Guidelines

Accurate diagnosis is a fundamental aspect of appropriate health care. However, clinicians need guidance when implementing diagnostic tests given the number of tests available and resource constraints in health care. Practitioners of health often feel compelled to implement recommendations in guidelines, including recommendations about the use of diagnostic tests. However, the understanding about diagnostic tests by guideline panels and the methodology for developing recommendations is far from completely explored (Hsu et al. 2011).

Therefore, Hsu et al. (2011) evaluated the factors that guideline developers and users need to consider for the development of implementable recommendations about diagnostic tests. The application of the GRADE approach to the development of diagnostic recommendations was described. This case study provides useful guidance for guideline developers and clinicians about what they ought to demand from clinical practice guidelines to facilitate implementation and strengthen confidence in recommendations about diagnostic tests. The particular challenges of making diagnostic recommendations were met by developing evidence profiles explicitly defining the patient important outcomes associated with being classified as true positive, true negative, false positive, and false negative in addition to providing sensitivity and specificity of diagnostic tests that would determine the number of patients that would fall into each group. Furthermore, there was an emphasis on diagnosis concepts in evidence-based medicine, such as a consideration of pretest probability and test/treatment thresholds. Throughout the guideline, a structured multidisciplinary panel process ensured the methodological soundness and clinical relevance of recommendations (Hsu et al. 2011).

Table 13.8 A proposed reporting checklist of systematic review or meta-analysis of diagnostic accuracy studies (Manchikanti et al. 2009b) (Reprinted with permission from the American Society of Interventional Pain Physicians)

Heading	Subheading	Descriptor
Title		Identify the report as a systematic review or meta-analysis of diagnostic accuracy studies.
Abstract		The abstract must utilize a structured format.
	Background	Problem definition Hypothesis statement
	Study Design	Clearly specify the nature of the design, either it is systematic review or meta-analysis or both.
	Objective(s)	Identify the objective(s) of the systematic review.
	Methods	Identify the methods related to literature search, inclusion criteria, method of review or validity assessment.
	Level of Evidence	Identify the nature of determination of level of evidence.
	Outcome	Measures Identify the outcome measures utilized in the accuracy studies.
	Results	Identify characteristics of the diagnostic accuracy studies included and excluded; qualitative and quantitative findings; and subgroup analysis.
	Limitations	Identify the limitations of the systematic review.
	Conclusion	Identify the main results.
Introduction		Describe the explicit clinical problem, biological rationale for the intervention, and rationale for the review
Methods	Literature Search	Provide description of comprehensive literature search with databases, registers, personal files, expert informants, agencies, hand searching and any restrictions (years considered, publication status, language of publications).
	Selection Criteria	Describe the inclusion and exclusion criteria with definition of the population, intervention, and outcomes.
	Method of review or validity assessment	Describe the methodologic quality assessment and the instrument utilized to achieve such an assessment.
	Data extraction	Describe the process of data extraction and the process utilized (e.g., completed independently, in duplicate, or other methodology utilized).
	Study characteristics	Describe the type of study design, participants' characteristics, details of intervention, outcome definitions, etc., and how clinical heterogeneity was assessed Describe the reference standard.
	Quantitative data synthesis	Describe validity, assessment bias and variation, synthesis of relevant study results, descriptive or non-qualitative synthesis, and quantitative synthesis or meta-analysis.
	Analysis of Evidence	Provide the results of analysis of evidence with strength and level of evidence and if applicable, with grading of recommendations.
Results	Trial flow	Provide a systematic review or meta-analysis profile summarizing the trial flow.
	Study characteristics	Present descriptive data of methodological quality assessment and diagnostic accuracy, including prevalence, confounding factors, study designs, and criterion standard.
	Quantitative data synthesis	Report agreement on the selection and validity assessment with presentation of simple summary results with statistical analysis descriptions.
Discussion		Summarize key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (e.g., publication bias, selection bias, etc.); describe limitations and weaknesses; and suggest a future research agenda.

In summary, applying a structured framework like the GRADE approach with its requirement for transparency in the description of the evidence and factors that influence recommendations facilitates laying out the process and decision factors that are required for the development, interpretation, and implementation of recommendations about diagnostic tests (Hsu et al. 2011).

References

- Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ*. 1994a;309:102.
- Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. *BMJ*. 1994b;309:188.
- Altman DG, Bland JM. Diagnostic tests. 1: sensitivity and specificity. *BMJ*. 1994c;308(6943):1552.
- Andermann A, Blancaquaert I, Beauchamp S, Déry V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bull World Health Organ*. 2008;86:241–320.
- Armitage GC. Periodontal diagnoses and classification of periodontal diseases. *Periodontol* 2000. 2004;34:9–21.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003a;326:41–4.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG. The STARD Statement for reporting studies of diagnostic accuracy. Explanation and elaboration. *Clin Chem*. 2003b;49:7–18.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC. Standards for reporting of diagnostic accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003c;326(7379):41–4.
- Deeks JJ. Using evaluations of diagnostic tests: understanding their limitations and making the most of available evidence. *Ann Oncol*. 1999;10:761–8.
- Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ*. 2004;329(7458):168–9.
- Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, Bezemer PD. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol*. 2002;2:9.
- Halligan S. Systematic reviews and meta-analysis of diagnostic tests. *Clin Radiol*. 2005;60:977–9.
- Hsu J, Brożek JL, Terracciano L, Kreis J, Compalati E, Stein AT, Fiocchi A, Schünemann HJ. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implement Sci*. 2011;6:62.
- Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, Mosteller F. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med*. 1994;120:667–76.
- Jaesche R, Guyatt GH, Sackett DL. Users' guides to the medical literature, VI. How to use an article about a diagnostic test. *JAMA*. 1994;271:389–91.
- Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149:889–97.
- Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JHP, Bossuyt PM. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. 1999;282:1061–6.
- Manchikanti L, Derby R, Wolfer L, Singh V, Datta S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 5. Diagnostic accuracy studies. *Pain Physician*. 2009a;12: 517–40.
- Manchikanti L, Derby R, Wolfer L, Singh V, Datta S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 7: systematic reviews and meta-analyses of diagnostic accuracy studies. *Pain Physician*. 2009b;12:929–63.
- Mandel ID. Overview of clinical trials of periodontal diagnosis methods and devices. *Ann Periodontol*. 1997;2:98–107.
- Needleman I, Moles DR. A guide to decision making in evidence-based diagnostics. *Periodontol* 2000. 2005; 39:164–77.
- Nirula R, Brasel K. Determining and interpreting the accuracy of a test. *Curr Surg*. 2006;63:190–3.
- Sousa MR, Ribeiro AL. Systematic review and meta-analysis of diagnostic and prognostic studies: a tutorial. *Arq Bras Cardiol*. 2009;92(3):229–39, 235–45.
- Streiner DL, Norman GR. "Precision" and "accuracy": two terms that are neither. *J Clin Epidemiol*. 2006;59: 327–30.
- Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003; 3:25.
- Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med*. 2004;140:189–202.
- Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, The QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–36.

Clinical Trials Principles for Evaluation of Antimicrobial Drugs in Periodontal Disease Treatment

14

According to the classic version of the non-specific theory in microbial aetiology of inflammatory periodontal disease, the indigenous oral bacteria in the absence of oral hygiene colonize the gingival crevice to form plaque. Inflammatory periodontal disease develops in the case of bacterial proliferation above the threshold of host resistance, caused by the combined biologic effects of the total plaque flora. All plaque bacteria are thought to have some of the various virulence factors causing gingival inflammation and periodontal destruction, and it is implied that plaque will cause disease regardless of its composition. Therefore, total plaque control is considered necessary in treatment and prevention, and this approach has certainly proved effective, where it can be carried out (Theilade 1986).

Some terms have previously been defined by the 2nd European Workshop on Periodontology in 1996 (Addy and Moran 1997, 2007):

- Antimicrobial agents: chemicals that have a bacteriostatic or bactericidal effect *in vitro* that alone cannot be extrapolated to a proven efficacy *in vivo* against plaque
- Plaque reducing/inhibitory agents: chemicals that have only been shown to reduce the quantity and/or affect the quality of plaque, which may or may not be sufficient to influence gingivitis and/or caries
- Antiplaque agents: chemicals that have an effect on plaque sufficient to benefit gingivitis and/or caries
- Antigingivitis agents: chemicals which reduce gingival inflammation without necessarily

influencing bacterial plaque (includes anti-inflammatory agents)

Among clinical investigations in periodontology, clinical studies that test the effectiveness of chemotherapeutic agents are of special interest. These studies are aimed to test agents that inhibit plaque formation, influence plaque removal, and prevent or reduce gingivitis and calculus (Lorenz et al. 2009a). The design of these studies was a matter of discussion in the past in order to find consensus about important details of the protocol that can affect the outcome (Chilton 1992; Fischman 1979; Newman 1997). Agreement was established in many areas and guidelines could be accomplished (Council on Dental Therapeutics 1986; Imrey et al. 1994a, b; ADA Council on Scientific Affairs 2007, 2008).

Clearly, clinical significance is in the eye of the beholder, be that a generalist, specialist, patient, or regulatory agency, and is based on the difference in outcome between control and test groups and the standard deviation of the measurement of each group. To determine clinical significance, the overall purpose of the study, its design, and the size of the effect that may matter to a patient must first be determined (Jeffcoat 2002).

Because this is a symposium from the 2001 American Association for Dental Research meeting, it is worthwhile to review the types of studies in US Food and Drug Administration (FDA) terminology. A phase I study is intended to demonstrate

safety, usually in healthy patients. In phase I studies, efficacy is not a major factor. Clinical significance is not an issue (Jeffcoat 2002).

Phase II studies are pilot studies intended to demonstrate safety and efficacy in a relatively small number of patients with periodontitis. In phase II studies, the sponsor may elect to use especially sensitive methods that will give an indication of the efficacy of a new therapy in a short period of time. Examples would include digital or subtraction radiology and measures of mediators. Alternatively, classic measurements, such as probing attachment levels, may be used as the major outcome (Jeffcoat 2002).

Phase III studies are often called pivotal studies, designed to assess safety and efficacy in larger numbers of subjects with disease. In a pivotal study, the nature of the population must be clearly defined and should reflect the target population that is to ultimately receive the treatment. Outcomes for the study (such as probing depths, clinical attachment levels, bleeding on probing, and/or bone levels) must be defined prior to the start of the study. The FDA requires at least two well-controlled clinical trials for the approval of a drug. In clinical practice, the definition of success may be modified over the course of treatment during re-evaluation visits. In a clinical trial, the criteria for success should be predetermined prior to the start of the study. Safety will be determined by comparing the nature and frequency of adverse effects in the treated group with that in the control group (Jeffcoat 2002).

Phase IV is the post-market surveillance phase. As more patients are treated during the period of marketing, this is an excellent opportunity to determine the incidence of rare side effects (Jeffcoat 2002).

14.1 Studies *In Vitro*

In order to test the antimicrobial activity of anti-septic products and to identify their possible spectrum of action, different *in vitro* tests have been used. Although conclusions from this type of studies cannot be drawn for *in vivo* conditions,

they are considered powerful tools for the purpose of comparison (Herrera et al. 2003).

14.1.1 Bacterial Tests

Antimicrobial tests, including minimum inhibitory concentration (MIC), minimum bactericidal concentration (MBC), and kill curves, can be determined. These tests indicate that antibacterial activity and antimicrobial spectrum of agents and formulations against a range of bacteria. Continuous culture technique can also be used, but they may not provide more meaningful data. At present, antimicrobial tests *in vitro* primarily only indicate activity, or lack of it, and they are poor predictors per se of effects on plaque *in vivo* (Addy and Moran 2007).

14.1.2 Uptake Measurements

One aspect of substantivity is adsorption of antimicrobials and other potential plaque-inhibitory agents on to surfaces. This can be quantified using a variety of substrates such as hydroxyapatite, dentin, enamel, acrylic, and other polymers. Again, data seem to be unreliable predictors of action *in vivo* since they only measure uptake, not activity once adsorbed. Obvious applications would be in the evaluation of antiadhesives and plaque removal agents but, for example, with the former, findings are usually more encouraging *in vitro* than *in vivo*. These experiments provide some information on alterations in surface free energy of dental structures produced by chemicals. Such data are relevant mainly to antiadhesive agents but alone correlate poorly with effects *in vivo*. In particular, such methods, even performed *in vivo*, cannot allow for the duration of action of the agent nor the versatility of primary plaque-forming bacteria (Addy and Moran 1997).

14.1.3 Other Methods

A number of *in vitro* tests have been developed that are peripheral to preventive or therapeutic oral health benefits of products. These are largely associated with cosmetic effects, such as stain

removal, or detrimental actions, such as abrasivity. Both are interrelated since scratching and polishing are aspects of abrasion. Methods *in vitro* have therefore evolved to measure the removal of stain from surfaces, both as a function of mechanical, chemical, or both actions. Abrasivity of toothpaste has been measured by a variety of methods including profilometry or surfometry, gravimetry, scanning electron microscopy, chemical analysis, and radiometry. A number of substrates have been used notably enamel, dentin, and acrylic. Similar methods can also be used to study the potential of some oral hygiene products to erode enamel and/or dentin. Nevertheless, tests *in vivo*, particularly of abrasion and erosion, are difficult to conduct, although modelling with inserts has been attempted. Stain control is perhaps more easily modelled, particularly using forced staining methods involving chlorhexidine (Addy and Moran 1997 and references herein).

14.2 Studies *In Vivo*

A wide range of clinical studies has been presented over the years in support of the effectiveness of therapeutic antimicrobial mouthrinses. At the least rigorous end of the spectrum are open trials, or so-called pragmatic clinical trials, in which a test group of subjects uses a product according to label directions with full knowledge of what the product is. In these trials, there is no true control group, as the test group is compared with a group of people who have done nothing in addition to their usual oral hygiene procedures. Since it is known which product the test subjects used, such a design does little to minimize investigator bias and in the absence of a true negative control group, does not allow for the factoring out of improvement, resulting from a 'Hawthorne effect'. Trials using this protocol design have only limited value in determining the effectiveness of a given mouthrinse formulation (Barnett 2003).

In general, there have been two sets of clinical trials of antiplaque agents—those involving 'restricted oral hygiene' or 'no-brushing' studies and those permitting normal oral hygiene practices by the participants. The '**no-brushing**' studies

followed the now-classic procedure reported by Loe et al. (1965) to demonstrate experimental gingivitis in man. Various reasons justify this no-brushing regimen. Primarily, the studies have been designed to screen the gingivitis-preventing potential of known plaque inhibitors. The restriction of normal oral hygiene fosters the rapid development of gingivitis and permits a shorter and more easily managed clinical trial. Further, the use of a dentifrice in clinical trials of mouthrinses creates questions of chemical compatibility. For example, the ionic detergents present in dentifrices have the ability to bind quaternary ammonium substances and thus decrease their bioavailability. Since any active agent would be recommended for use in a brushing population, it must be demonstrated to be effective in a population practicing at least normal oral hygiene. Many agents shown to be effective gingivitis inhibitors in no-brushing studies have not been able to produce statistically or clinically significant reductions in gingival disease when used by brushing subjects (Fischman 1979).

Another variable is **the use of a dental prophylaxis (scaling and polishing) prior to the study**. Plaque studies have usually been done on subjects rendered plaque-free at the start of the study by a thorough polishing of the teeth. They have measured the inhibition of plaque formation, rather than the removal of pre-existing plaque. In a clinical sense, then, these agents would likely be prescribed by the dentist for use following a thorough professional prophylaxis. They would not be dispensed *ad libitum* to the general population, most of whom receive no regular professional care (Fischman 1979).

There are four characteristics that are essential in assessing which types of clinical studies are the most credible (Barnett 2003):

- The trial should be controlled with a negative control group to which the test group can be compared.
- The trial should be blinded so that the examiner and all other personnel involved in evaluating data or laboratory specimens are not aware of which group a given subject is in.
- Subjects meeting entry criteria should be assigned randomly to treatment and control groups.

- The trial should be prospective; that is, the protocol, entry criteria, random assignments, and method of statistical analyses should be determined in advance of the actual treatment phase.

14.2.1 The 8-h Substantivity Studies

The design of 8-h substantivity studies was first used and described by Bonesvoll et al. (1974) and afterwards applied in many studies (Addy et al. 1989, 1991; Bonesvoll and Olsen 1974; Bonesvoll 1977; Gjermo et al. 1974; Jenkins et al. 1990; Fine et al. 2005; van der Mei et al. 2006; Lorenz et al. 2009c). As **short-term test**, the aim of substantivity studies is to determine whether or not and if for how long a formulation performs a persisting effect *in vivo*. This substantivity depends on the ability of a substance of physical and chemical bonding to a surface as well as its resistance against removal or inactivation, as long as it remains biologically active (Lorenz et al. 2009a). As the first test stage, failure of a formulation to show substantivity would prove an inappropriate effect on the inhibition of plaque development and her ability to reduce the proportion of vital bacteria in established plaque biofilm following a single rinsing. If there is an effect then according to the obtained efficacy profile, an optimal frequency of use for any product can be recommended. The substantivity of an agent determines the rinsing frequency needed; however, practically, the rinsing frequency is limited to two or three times a day. After application of a single rinse on pre-existing plaque, plaque and saliva bacteria are studied for the following 8 h, while the participants cease oral hygiene measures (Lorenz et al. 2009a).

14.2.2 Experimental Plaque Studies

The 24-h plaque regrowth model proposed by Harrap (1974) has been used successfully in several studies by the present group (Claydon and Addy 1995, 1999; Claydon et al. 1996, 2002). The model is aimed to study the plaque-inhibitory effect of a formulation *in vivo*, while any oral hygiene is stopped during the test phase (Lorenz et al. 2009a). The procedure, as described by Claydon et al. (2002) is as follows: On day 1 of the study period, subjects received a prophylaxis to remove all plaque, supragingival calculus and extrinsic stain from the teeth. Oral hygiene practices were then suspended and subjects commenced the rinsing regimen with the allocated mouth rinse. After 24 h, subjects returned to the clinic where, after disclosing the accumulated plaque with erythrosine, plaque was scored by a single examiner (Claydon et al. 2002).

The 4-day plaque regrowth model was described by Addy et al. (1983): Subjects received a session of oral prophylaxis to remove any plaque, calculus, and stain and were given a supply of the allocated formulation. Then, the subjects were instructed to withdraw from any oral hygiene measures for the following 4 days. During this period, subjects were instructed to follow their normal diet, to avoid the use of chewing gum, and to use the allocated products. The assigned mouthrinse regime was the only oral hygiene procedure allowed during the experimental period. On day 5, plaque was scored and smears were collected according to the protocol (Moran et al. 1994; Herrera et al. 2005; Pretty et al. 2004; Cosyn et al. 2005; Claydon et al. 2005; Welk et al. 2005; Arweiler et al. 2006; Rodrigues et al. 2010) (Fig. 14.1). A final plaque

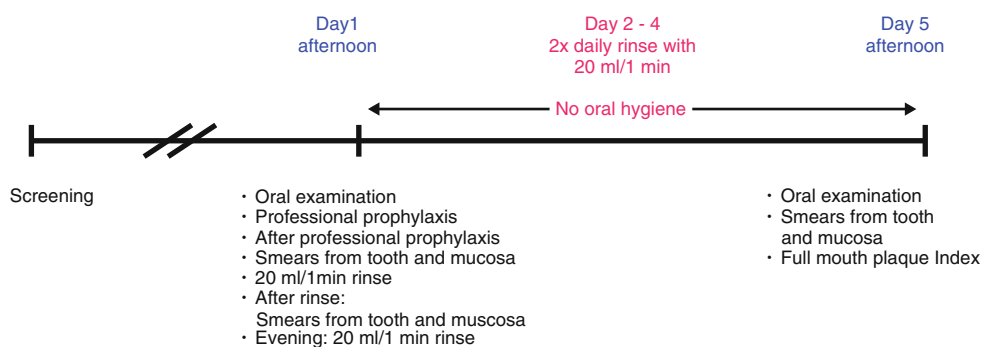


Fig. 14.1 Outline of the clinical trial (Welk et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

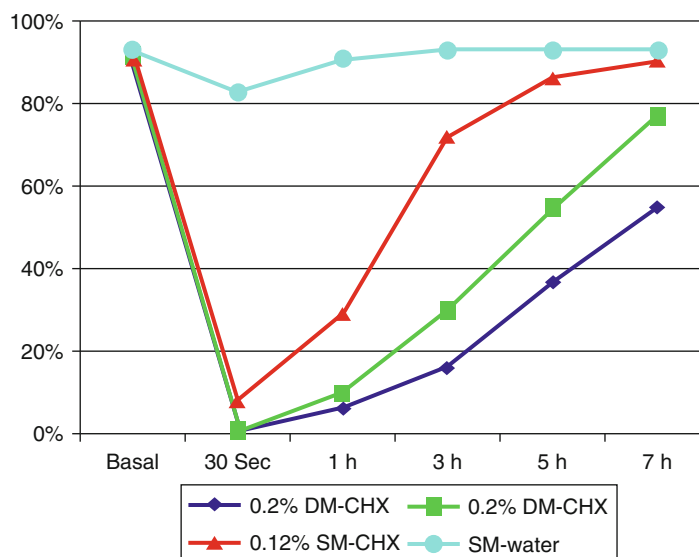


Fig. 14.2 Percentages of bacterial vitality in saliva under basal conditions and at 30 s and 1, 3, 5, and 7 h after the application of sterile water (single mouthrinse); 0.12% chlorhexidine (single mouthrinse); and 0.2% chlorhexidine (single and double mouthrinse). *SM* single mouthrinse, *DM* double mouthrinse, *CHX* chlorhexidine. Basal saliva sample collected under basal conditions, 30-s saliva sample collected at 30 s after the application of the

different mouthrinses, 1-h saliva sample collected 1 h after the application of the different mouthrinses, 3-h saliva sample collected 3 h after the application of the different mouthrinses, 5-h saliva sample collected 5 h after the application of the different mouthrinses, 7-h saliva sample collected 7 h after the application of the different mouthrinses (Cousido et al. 2010. Reprinted with permission from Springer Science+Business Media)

assessment shows whether or not the mouthrinse per se is able to depress plaque development and to what extent. If no plaque inhibition can be shown in this type of study, no further effect of the rinsing solution can be expected in studies when oral hygiene is performed (Lorenz et al. 2009a).

14.2.3 Bacteriology *Ex Vivo*

Substances are used in the oral cavity, that is, *in vivo*, while the bacteriological samples (saliva or the dental biofilm) are taken out of the mouth to be evaluated '*ex vivo*' in the laboratory (Netuschil et al. 2003).

One of the most accepted designs is the evaluation of the **salivary bacterial counts** after a single use of the tested product, in a crossover design (Fig. 14.2) (Elworthy et al. 1996; Addy et al. 1997).

The first postrinsing nonstimulated saliva sample is usually taken after 30 min (Addy et al. 1991, 1997; Jenkins et al. 1991; Elworthy et al. 1996); however, slight modifications can be found in different publications (Yates et al. 1997;

Fine et al. 2000; Herrera et al. 2003; de Albuquerque et al. 2004; Cousido et al. 2008, 2010; Botelho et al. 2009).

The bacteriological methodology may also differ, and this might also influence the results. Most of the studies have only assessed anaerobic bacteria (Addy et al. 1991, 1997; Jenkins et al. 1991; Harper et al. 1995; Moran et al. 1995; Elworthy et al. 1996); however, the addition of aerobic bacteria can provide valuable information to the results, as shown in the present and other studies (Moran et al. 1988). The time of incubation also shows some variability. One study employed an anaerobic incubation period of 12 h (Addy et al. 1997), while others extended that period up to 72 h (Elworthy et al. 1996; Harper et al. 1995; Cousido et al. 2008). The most frequent incubation period is 48 h (Moran et al. 1988, 1995; Jenkins et al. 1990; Addy et al. 1991; Herrera et al. 2003; Botelho et al. 2009). For aerobic incubation, 24 h were used (Moran et al. 1988; Herrera et al. 2003; Cousido et al. 2008).

The procedure measures both the antibacterial action *in vivo* and the persistence of this effect

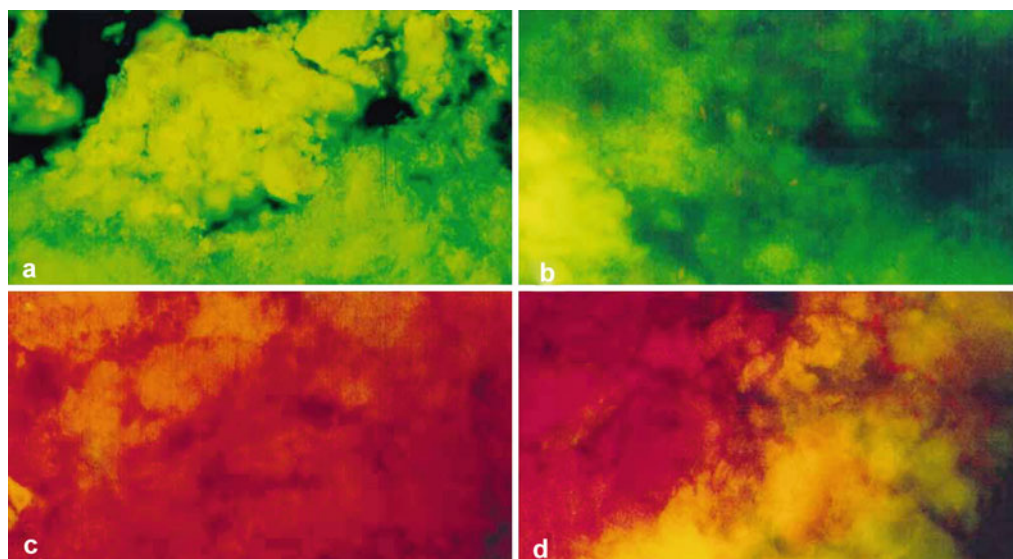


Fig. 14.3 (a) Photomicrograph of baseline vital stained plaque sample. Predominantly, *green staining* indicates live bacteria (orig. mag. $\times 25$). (b) Photomicrograph of control plaque sample at 30 min (orig. mag. $\times 25$). (c, d)

Photomicrographs of 30-min plaque samples from essential oil mouthrinse group showing predominantly *red staining* (d) and complete *red staining* (c) (orig. mag. $\times 25$) (Pan et al. 2000. Reprinted with permission John Wiley and Sons)

(Jenkins et al. 1994). Therefore, it is considered a useful test to assess substantivity (Schiott et al. 1970) and antiplaque activity (Roberts and Addy 1981). However, the salivary counts reflect the planktonic phase of the oral bacteria, while the true target of mouthrinses is the adherent dental biofilm, previously called dental plaque. It is important to note that biofilm microorganisms are much more resistant against antibiotics and antiseptic than planktonic bacteria, and therefore, the results have to be interpreted carefully (Netuschil et al. 2003). Only when the concentration of active ingredients in mouthrinse preparations reaches about 100 times the MBC, an antibacterial action on the biofilm microbiota is to be expected (Netuschil et al. 2003; Socransky and Haffajee 2002).

Vital fluorescence is very suitable for biofilm research (Pan et al. 2000). Vital fluorescence technique renders the vitality pattern of bacterial plaque visible, thus providing a quantitative estimation of living and dead plaque, whereas indices measure total plaque amounts only, that is, regardless of cell vitality. This technique quantifies the antiplaque effect due to bactericidal activity without

taking into account other antiplaque mechanisms such as antiadhesiveness (Fig. 14.3) (Weiger et al. 1998; König et al. 2002).

This technique was used for the assessment of the vitality of biofilm microorganisms throughout mouthwash application or toothpaste slurries in short-term investigations as well as in experimental gingivitis and long-term studies (Arweiler et al. 2000, 2001a, b, 2002a, b, 2003, 2006, 2008; Auschill et al. 2001, 2005; Brex et al. 1990, 1993; Gehlen et al. 2000; König et al. 2002; Lorenz et al. 2009; McBain et al. 2003; Netuschil et al. 1989, 1995, 1996, 1998; Rundegren et al. 1992; von Ohler et al. 1998; Weiger et al. 1995, 1998). For example, 1 min of rinsing with 0.2% chlorhexidine reduced the vitality rate by 21.82% with cold solution and by 47.21% with warm solution (König et al. 2002). Following two rinsing procedures with nonheated 0.1% chlorhexidine, Netuschil et al. (1989) observed a reduction of vitality of 57% in 3-day-old plaque.

The vital fluorescence is an additional tool available, which is not meant to take the place of other indices but rather to elucidate the mechanisms of action of mouthrinses. It does not give any

information about plaque accumulation, as many dead bacteria are present in the biofilm already from the start of its formation (Netuschil et al. 1996, 2003).

14.2.4 Experimental Gingivitis Studies

Experimental gingivitis studies over *are short-term 3-week studies* and prove the etiological role of plaque in the development of gingivitis. A shorter investigation time less than 21 days was not taken into consideration because the original studies of L  e et al. (1965) had already shown that some individuals need a longer time than 14 days to develop plaque-induced gingivitis. During 21 days, the participants stop their habitual oral hygiene. Instead, they rinse with the allocated mouthrinse once or twice a day. The ability of a formulation to inhibit gingivitis and plaque is assessed weekly. Once a test agent has shown its potential to inhibit plaque, this model elucidates whether the plaque-inhibiting effect also affects the development of gingivitis (Lorenz et al. 2009a).

An important modification of the **full-mouth model** has been the use of **partial-mouth experimentation** (two quadrants, one quadrant, three teeth, and any number of contiguous teeth extending over or two quadrants) (Bosman and Powell

1977; Matheny et al. 1993; Putt et al. 1993; van der Weijden et al. 1994; Daly and Highfield 1996; Deinzer et al. 1999; Saxton and van der Ouderaa 1989; Nogueira-Filho et al. 2000; Grant et al. 2010) as opposed to total oral hygiene abstinence. The model involves periodontally and systemically healthy volunteers, who do not brush specified test teeth in one sextant of the dentition, but continue to brush the remaining control teeth over a 21-day period. Accidental brushing of test teeth is prevented by the use of a vinyl shield (mouthguard) to cover them during cleaning. On day 21, the accumulated plaque on all teeth is removed by a trained hygienist, and study participants resume their normal oral hygiene practices. Clinical assessment of plaque accumulation and gingival inflammation are made at baseline, day 7, 14, and 21 on test and control teeth, following plaque accumulation on test teeth, and normal plaque removal from control teeth. Assessments are repeated again on day 35, following 14 days of resolution of inflammation at test sites, and maintenance of health at control sites. When performed by trained and calibrated experts, these measures show consistent increases in plaque-induced inflammation up to day 21 followed by a return to baseline, preinflammation levels at day 35 (Fig. 14.4) (Grant et al. 2010). Although the ‘partial-mouth’ assessment may have potentially

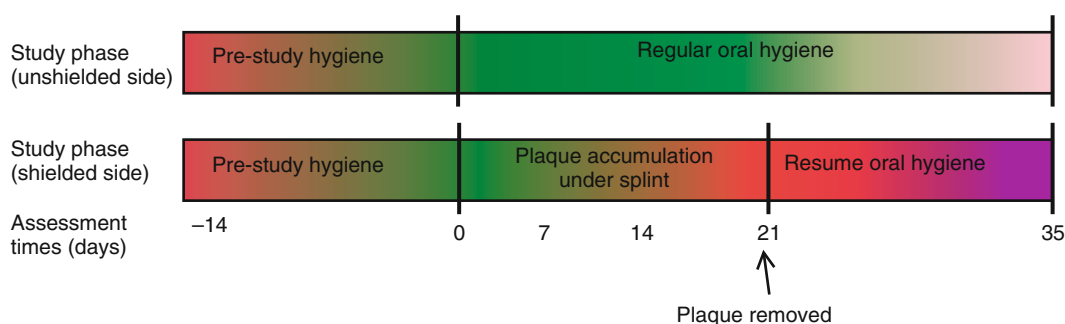


Fig. 14.4 Flowchart of study design for test sites. Control sites were treated identically until baseline but normal plaque removal then continued until day 35. A split mouth design was employed with test sites being the maxillary left 4, 5, and 6 (first and second premolars and first molar on the left-hand side of the upper jaw) in right-handed individuals and the equivalent teeth on the maxillary right side in left-handed individuals. Control sites were the corresponding teeth on the contralateral side. A soft vinyl splint was

constructed to cover five teeth, with a 5-mm clearance at each free surface from the marginal tissues. The splint was inserted gently over test teeth before brushing, to ensure against any mechanical or chemical cleaning. The test and control teeth were given a thorough prophylaxis, and oral hygiene was closely monitored for 2 weeks prior to the study, commencing to ensure pristine gingival health at baseline (Reprinted with permission from Grant et al. (2010). Copyright 2012 American Chemical Society)

limited the available information on the extent and rate of plaque accumulation and gingival inflammation compared with the 'full-mouth' assessment, this modification made the experimental gingivitis model easier to accept by the prospective participants, therefore more feasible from a practical perspective. It also lessened the probability of adverse effects by limiting the number of teeth on which plaque was allowed to accumulate, rendering the model preferable from an ethical perspective. The results of partial-mouth ('localized') experimental gingivitis studies have always been similar to the results of full-mouth studies, making the 'localized' assessment equivalent to the original full-mouth one (Trombelli et al. 2008).

The experimental gingivitis model protocol requires subjects to reach clinical health before beginning the 21-day non-brushing period; however, the tendency of the subject to develop gingivitis may also have an impact on gingivitis outcome variables (Shearer et al. 2005). Evidence suggested early on that the onset, and severity of the gingival inflammatory response to plaque accumulation might differ significantly among individuals, with such differences essentially ascribed to differences in plaque accumulation rates (quantitative plaque differences) and/or differences in plaque species present (qualitative plaque differences) (Löe et al. 1965; Trombelli et al. 2006). The reported significant differences in gingival inflammatory response under quantitatively and/or qualitatively almost identical plaque accumulation (Trombelli et al. 2004a) suggest that the level of the gingival tissue response to plaque accumulation may be an individual trait (Tatakis and Trombelli 2004; Trombelli et al. 2006). An immediate implication of such a tenet is that a subject's gingival inflammatory response will be consistently high or low relative to the level of plaque exposure (Trombelli et al. 2006); where the identified '**high-responder**' group—when compared with the '**low-responder**' group—had a significantly greater gingival inflammatory response after either 7 or 14 days of de novo plaque accumulation (Trombelli et al. 2004a, b), significantly higher GCF levels even in areas of the

dentition where ideal plaque control was maintained (Trombelli et al. 2004a, c), and significantly higher GCF levels in areas of the dentition where ideal plaque control was re-established after implementation of a therapeutic regimen (Trombelli et al. 2004c). McClanahan and Bartizek (2002) reported that subjects with higher baseline gingival bleeding scores were more likely to show reductions in gingival bleeding and gingival severity index using triclosan/copolymer in a 3-month clinical trial.

In experimental gingivitis studies, while plaque scores remained similar between smokers and non-smokers, the bleeding scores were half that in the smokers compared with non-smokers even though no major differences in the oral microbiota could be determined (Lie et al. 1998a, b). Considering the detrimental effect of smoking on periodontal tissues, there is sufficient evidence to justify **balancing for (or excluding) smokers** from subject populations in experimental gingivitis studies (Shearer et al. 2005).

Characteristic bacteriological changes were revealed in the plaque along the gingival margin during this experiment. After 9–21 days without oral hygiene, 11 experimental subjects with previously excellent oral hygiene and healthy gingivae developed heavy accumulations of plaque and generalized mild gingivitis. Initially, that is, when the teeth were clean and the gingiva healthy, the extremely sparse plaque flora consisted almost exclusively of gram-positive cocci and rods. The first phase of plaque development occurred during the first 2 days without oral hygiene and consisted of a proliferation of the gram-positive cocci and rods and an addition of about 30% gram-negative cocci and rods. During the second phase, (after 1–4 days) fusobacteria and filaments appeared and increased until they each made up about 7% of the flora. During the third phase (after 4–9 days), the flora was supplemented with spirilla and spirochetes, and at the end of the period without oral hygiene, each of these two groups of organisms accounted for about 2% of the plaque flora. In specific areas, the gingival condition was correlated with the composition of the plaque, and it was found that mild gingivitis could be diagnosed clinically at approximately the same time as the

complex flora was established. However, subclinical inflammation started much earlier, probably as a reaction to the first phases of plaque development. When oral hygiene was reinstituted, the plaque in most areas disappeared in 1–2 days, and after 7–11 days, the plaque index for each subject was as low as before the experiment. Correspondingly, after 1–2 days, most tooth surfaces only harboured the original sparse flora of gram-positive cocci and rods. The gingival inflammation in an area usually disappeared one day after the plaque had been removed (Theilade 1966).

It has been showed that during a repeated experimental gingivitis trial, under well-controlled experimental conditions (inclusion of control quadrants, use of stents to avoid inadvertent plaque removal at test quadrants, vitamin C supplementation to reduce environmental effects of potential differences in vitamin C intake), the model is to some extent reproducible in selected populations. Specifically, the plaque accumulation parameters and the bleeding index values showed no differences, despite the time separating the trials and regardless of the group examined (Trombelli et al. 2008).

Although experimental gingivitis and **persistent gingivitis** reflect the host's response to plaque bacteria, they differ in at least three major aspects. First, experimental gingivitis is preceded by a prophylactic intervention, which might affect the host's response to subsequently accumulating plaque bacteria. Second, during experimental gingivitis, all oral hygiene procedures are abolished at the sites under study. Patients suffering from persistent gingivitis, in contrast, do show at least some oral hygiene behaviour; even though this behaviour is insufficient with respect to plaque removal, it still differs from complete neglect of oral hygiene. Third, under experimental gingivitis, plaque is not allowed to accumulate for more than 28 days. To avoid irreversible damage, oral hygiene has to be re-established after this time (Deinzer et al. 2007).

When the two conditions were compared in a randomized controlled design, the data indicate that conditions observed after 4 weeks experimental gingivitis are not comparable with

persistent gingival inflammation in a naturalistic setting. After 4 weeks, subjects with experimental gingivitis showed significantly more plaque accumulation, higher interleukin-1 β , and lower interleukin-8 concentrations than subjects with persistent gingivitis. Whereas in experimental gingivitis, we observed considerable fluctuations in clinical and immunological parameters over the 4-weeks period, and persistent gingivitis was characterized by little fluctuation, indicating to be a type of immunological and clinical steady state (Figs. 14.5, 14.6, and 14.7) (Deinzer et al. 2007).

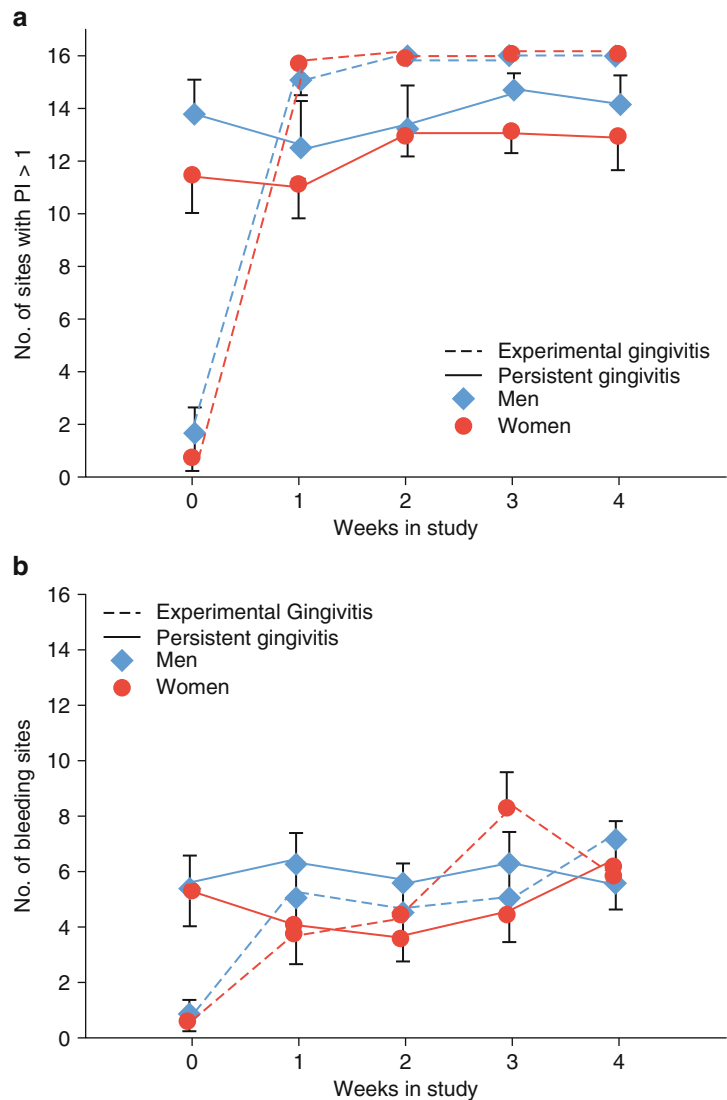
14.2.5 Home-Use Studies

Home-use studies are *6 months long-term studies* to test the efficacy of antiplaque and antigingivitis agents under almost real-life circumstances. This model refers to the FDA requirements that ask for safety records for oral hygiene products as well (Lorenz et al. 2009a). Clinical trials involving human patients, subjects, or healthy volunteers in many countries must conform to the Guidelines for Good Clinical Practice (ICH 1996) including the Declaration of Helsinki (World Medical Association 1996) (Addy and Moran 2007).

All studies evaluating the efficacy of an agent, formulation, or product should conform to scientific principles. As stated, the **blind, randomized, and suitably controlled clinical trial** is the basis from which to evaluate all formulations, including oral hygiene products (Addy and Moran 1997, 2007):

- **Blindness (Masking):** Single blindness requires that any investigator collecting data from the patients should not know the identity of the treatment used by any particular individual. Double blindness is the ideal when neither the examiner nor subject is aware of the treatment being used by the individual. The term triple blindness is used by some investigators and related to the blindness of the individual analysing the data to the identity of treatments.
- **The order in which treatments** are received by each subject in a crossover study, or into

Fig. 14.5 Plaque index and bleeding response in experimental and persistent gingivitis. Mean and standard error of the mean of the number of sites showing plaque indices above 1 (a), and a positive bleeding response (b), are shown for the 16 sites where gingival crevicular fluid samples were taken. Subjects in the experimental gingivitis group (six men, seven women) refrained from any oral hygiene for 4 weeks, whereas subjects in the persistent gingivitis group (seven men, six women) proceeded with their habitual oral hygiene behaviour. *PI* plaque index (Deinzer et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)



which treatment group subjects are placed in a parallel study, should be according to a randomization schedule.

- **The choice of controls** for clinical trials on oral hygiene could be one or more of the following:
 1. Placebo control: one without any expected activity against the condition under investigation
 2. Minus active ingredient control (negative control): one that is identical to the test product or formulation without the active agent or agents
 3. Benchmark control: a product already in use by the public or an agent already evaluated
 4. Positive control: an accepted effective formulation or the most effective formulation available to date

The ADA's Council on Dental Therapeutics (currently, the ADA Council on Scientific Affairs) developed guidelines for the design of such clinical trials (Council on Dental Therapeutics 1986). Pivotal studies included in product submissions for the **ADA Seal of Acceptance** are required to meet the requirements of these guidelines. These guidelines also have been adopted, in some cases with modifications, by other professional and governmental organizations that evaluate anti-plaque and antigingivitis products; these organizations include the US Food and Drug Administration, or FDA; the Canadian Dental Association; and the British Dental Association (Barnett 2003). Subsequent modifications to the

Fig. 14.6 Crevicular interleukin-1b in experimental and persistent gingivitis. Mean and standard error of the mean of standard curve units (ng/ml) are shown for the 16 sites under study. Subjects in the experimental gingivitis group (six men, seven women) refrained from any oral hygiene for 4 weeks, whereas subjects in the persistent gingivitis group (seven men, six women) proceeded with their habitual oral hygiene behaviour. IL-1b, interleukin-1 (Deinzer et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

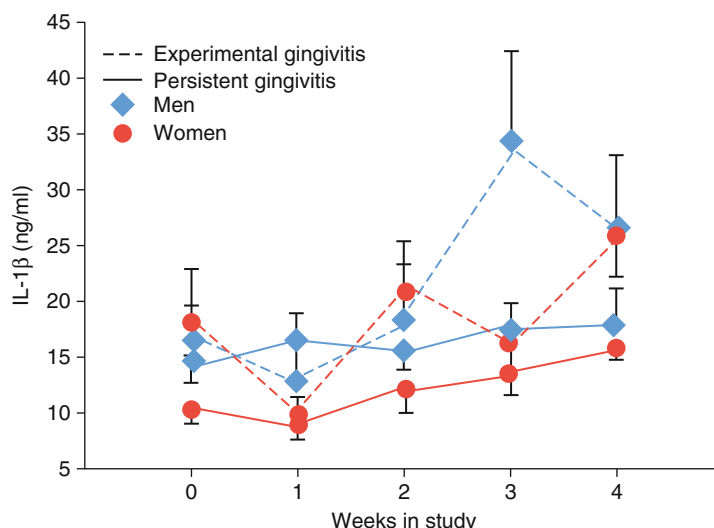
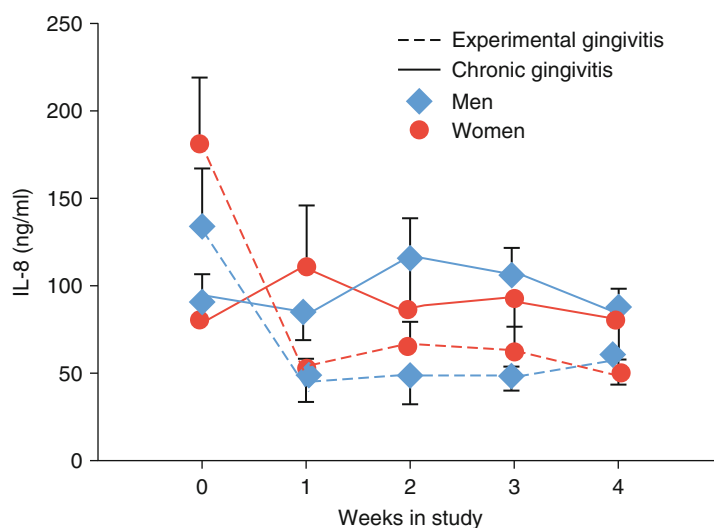


Fig. 14.7 Crevicular interleukin-8 in experimental and persistent gingivitis. Mean and standard error of the mean of standard curve units (ng/ml) are shown for the 16 sites under study. Subjects in the experimental gingivitis group (six men, seven women) refrained from any oral hygiene for 4 weeks, whereas subjects in the persistent gingivitis group (seven men, six women) proceeded with their habitual oral hygiene behaviour. IL-8, interleukin-8 (Deinzer et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)



guidelines (Imrey et al. 1994a, b) deal with issues of study design, such as making the need for randomization and blinding explicit; require a gingival bleeding component in the assessment of gingivitis; indicate methods for standardization of examiners; specify elements to be included in the statistical analyses; and establish a minimum acceptable effect level (Barnett 2003).

The recent Guidelines for Chemotherapeutic Products for Control of Gingivitis (The ADA Council on Scientific Affairs 2008) presents the clinical protocol guidelines given for the design and conduct of clinical studies for the evaluation of chemotherapeutic agents to provide evidence of effectiveness and safety in the control of gingi-

vititis and if applicable, supragingival plaque. In each study, the active product should be compared with a placebo control. In addition, a positive control may be added.

Other specifications refers to the:

- **Sample size:** A sufficient number of subjects should be enrolled in the study to ensure that appropriate statistical tests can be performed.
- **Study length:** The studies will be conducted for a minimum of 6 months.
- **Study design:** Studies should be conducted with populations and under conditions that the product is expected to be used. Both genders should normally participate, the age of the study populations and the levels of gingivitis



Fig. 14.8 Erythrocline stained teeth to record the plaque area after 4 days of rinsing with the following mouthwashes: (a) placebo, (b) amine fluoride/stannous fluoride

mouthrinse, (c) chlorhexidine (Arweiler et al. 2001a. Reprinted with permission from John Wiley & Sons, Inc.)

should be representative of those patients for whom the product is intended, the frequency of use of the product should be representative of actual use of the product in practice, and the user should be instructed in the proper use of the product but not necessarily supervised.

- Efficacy assessments include details regarding the plaque and gingivitis assessments as well as recommended statistical analysis (Fig. 14.8).
- Safety assessments of oral soft and hard tissue as well as microbial assessment of plaque are to determine whether there are shifts in the balance of flora that might have an adverse effect on oral tissues or contribute to the progression of microorganisms in the flora from non-periodontal pathogens to those that are pathogenic.

Additional information regarding clinical trials and clinical trial reporting can be obtained from the ADA Council's Guidelines for Clinical Trial Protocols (2007).

These studies are very labour intensive since they involve large numbers of subjects, often with a considerable number of indices and parameters measured at several time points during the study.

Additionally, where safety and side effects are monitored, in addition to microbiological assessments, a range of medical, biochemical, and haematological monitoring may be necessary. All of these increase the personnel, materials, costs, and complexity of the study (Addy and Moran 1997).

14.3 Class or Trial Type

14.3.1 Monotherapy Versus Adjunctive

One of the first clinical trial design decisions is whether the product will be used as the primary (monotherapy) or supportive (adjunctive) treatment. The potential market and use of a monotherapy drug device may be considerably larger than for an adjunctive product. The rationale for a monotherapy is also driven, in part, by cost-benefit, cost-effectiveness considerations. If these antimicrobial periodontitis drugs and devices can replace traditional mechanical intervention, extend the period of time between office visits,

and/or minimize the time each visit takes, there would be great benefits to patients, society, and manufacturers. Monotherapy products also have the potential for indiscriminate application and abuse. At the present time, a scenario where scaling, root planing, and debridement are not an essential part of overall treatment cannot be envisioned for periodontitis patients. In contrast to the monotherapy approach, a common claim strategy has been for adjunctive use. The adjunctive point of view states that antimicrobial periodontitis products must be shown to be effective by demonstrating healing of the periodontitis lesion over and above traditional therapy of scaling and root planing (Newman 1997).

14.3.2 Therapeutic or Preventive

The preceding discussion demonstrates that there is more than one philosophy of effective disease treatment, and it highlights the need to clarify and define exactly what are the acceptable clinical outcomes of treatment. Until recently, the major focus of periodontitis products was on the improving clinical status. Today, arrest of disease and prevention of progression are considered a sufficient outcome. Selecting one clinical trial goal over another depends upon many factors including the intended population, the intended use, superiority versus equivalency, and the exact nature of claim strategy (Newman 1997).

14.3.3 Superiority or Equivalency

In reporting clinical trials, the clinical significance of the findings should be considered. Clinical significance can be assessed as follows:

- Benchmark equivalent: when a formulation performs as well as an established formulation or product
- Benchmark superior: when a formulation performs significantly better than an established formulation
- Disease efficacy: achieving an effect on a causative factor that reduces a disease or condition
- Positive efficacy: achieving an effect significantly greater than the most effective formulation to date
- Proportional efficacy: achieving a previously agreed proportional reduction in a parameter compared with control (Addy and Moran 2007)

An important issue in evaluating the quality of RCTs is the distinction between superiority and equivalence. A clinical trial to test whether a new treatment modality has a genuine effect, or can give rise to better outcomes than the placebo or conventional treatment, is known as a superiority trial. A clinical trial to test whether the performance of a new treatment, which might be cheaper or easier to use, is comparable with that of the established treatment in current practice, is known as an equivalence trial (Tu et al. 2006).

An RCT using an established treatment as an active-control group can be a superiority trial, if the aim of the trial is to show that the new treatment is better than the established treatment. In contrast, an RCT using an established treatment as an active-control group can be an equivalence trial (or a non-inferiority trial), if the aim of the trial is to show that the new treatment is as good as (or at least as good as) the established treatment. However, the research hypotheses and study designs are entirely different for these two dissimilar research strategies. For instance, the required sample size for an equivalence trial might be substantially greater than that for a superiority trial (Tu et al. 2005). Underpowered active-controlled trials might give rise to a false impression that the new treatment is as good as the established one, if there is no distinction between superiority and equivalence in their research hypotheses (Tu et al. 2006).

References

- ADA Council on Scientific Affairs. Acceptance Program Guidelines for Clinical Trial Protocols. Chicago: American Dental Association; 2007. Retrieved on http://www.ada.org/sections/scienceAndResearch/pdfs/guide_clinical_trial.pdf.
- Addy M, Moran JM. Evaluation of oral hygiene products: science is true; don't be misled by the facts. *Periodontol* 2000. 1997;15:40–51.

- Addy M, Moran J. Chemical supragingival plaque control in periodontal practice. In: Lindhe J, editor. *Clinical periodontology and implantology*. Oxford: Blackwell; 2007. p. 734–65.
- Addy M, Willis L, Moran J. Effect of toothpaste rinses compared with chlorhexidine on plaque formation during a 4-day period. *J Clin Periodontol*. 1983;10:89–99.
- Addy M, Jenkins S, Newcombe R. Toothpastes containing 0.3% and 0.5% triclosan. II. Effects of single brushings on salivary bacterial counts. *Am J Dent*. 1989;2:215–9.
- Addy M, Jenkins S, Newcombe R. The effect of some chlorhexidine-containing mouthrinses on salivary bacterial counts. *J Clin Periodontol*. 1991;18:90–3.
- Addy M, Greenman J, Renton-Harper P, Newcombe R, Doherty F. Studies on stannous fluoride toothpaste and gel (2). Effects on salivary bacterial counts and plaque regrowth in vivo. *J Clin Periodontol*. 1997;24:86–91.
- Arweiler NB, Donos N, Netuschil L, Reich E, Sculean A. Clinical and antibacterial effect of tea tree oil-a pilot study. *Clin Oral Investig*. 2000;4:70–3.
- Arweiler NB, Netuschil L, Reich E. Alcohol-free mouthrinse solutions to reduce supragingival plaque regrowth and vitality. A controlled clinical study. *J Clin Periodontol*. 2001a;28:168–74.
- Arweiler NB, Reich E, Brex M, Dörmer M, Netuschil L. Influence of ethanol on the antibacterial and anti-plaque efficacy of an amine fluoride/stannous fluoride mouthrinse solution. *J Parodontol Implantol Orale*. 2001b;20:331–40.
- Arweiler NB, Auschill TM, Reich E, Netuschil L. Substantivity of toothpaste slurries and their effect on reestablishment of the dental biofilm. *J Clin Periodontol*. 2002;29:615–21.
- Arweiler NB, Henning G, Reich E, Netuschil L. Effect of an amine-fluoride-triclosan mouthrinse on plaque regrowth and biofilm vitality. *J Clin Periodontol*. 2002b;29:358–63.
- Arweiler NB, Auschill TM, Baguley N, Netuschil L, Sculean A. Efficacy of an amine fluoride-triclosan mouthrinse as compared to the individual active ingredients. *J Clin Periodontol*. 2003;30:192–6.
- Arweiler NB, Boehnke N, Sculean A, Hellwig E, Auschill TM. Differences in efficacy of two commercial 0.2% chlorhexidine mouthrinse solutions: a 4-day plaque regrowth study. *J Clin Periodontol*. 2006;33:334–9.
- Arweiler NB, Lenz R, Sculean A, Al-Ahmad A, Hellwig E, Auschill TM. Effect of food preservatives on in situ biofilm formation. *Clin Oral Investig*. 2008;12:203–8.
- Auschill TM, Arweiler NB, Netuschil L, Brex M, Reich E, Sculean A. Spatial distribution of vital and dead microorganisms in dental biofilms. *Arch Oral Biol*. 2001;46:471–6.
- Auschill TM, Hein N, Hellwig E, Follo M, Sculean A, Arweiler NB. Effect of two antimicrobial agents on early in situ biofilm formation. *J Clin Periodontol*. 2005;32:147–52.
- Barnett ML. The role of therapeutic antimicrobial mouthrinses in clinical practice: control of supragingival plaque and gingivitis. *J Am Dent Assoc*. 2003;134:699–704.
- Bonesvoll P. Influence of ionic strength, calcium, sodium dodecyl sulphate and urea on the retention of chlorhexidine in the human mouth after mouth rinses. *Arch Oral Biol*. 1977;22:273–9.
- Bonesvoll P, Olsen I. Influence of teeth, plaque and dentures on the retention of chlorhexidine in the human oral cavity. *J Clin Periodontol*. 1974;1:214–21.
- Bonesvoll P, Lokken P, Rolla G, Paus PN. Retention of chlorhexidine in the human oral cavity after mouth rinses. *Arch Oral Biol*. 1974;19:209–12.
- Bosman CW, Powell RN. The reversal of localized experimental gingivitis. A comparison between mechanical toothbrushing procedures and a 0.2% chlorhexidine mouthrinse. *J Clin Periodontol*. 1977;4:161–72.
- Botelho MA, dos Santos RA, Martins JG, Carvalho CO, Paz MC, Azenha C, Ruela RS, Queiroz DB, Ruela WS, Marinho G, Ruela FI. Comparative effect of an essential oil mouthrinse on plaque, gingivitis and salivary *Streptococcus mutans* levels: a double blind randomized study. *Phytother Res*. 2009;23:1214–9.
- Brex M, Netuschil L, Reichert B, Schreil G. Efficacy of Listerine, Meridol and chlorhexidine mouthrinses on plaque, gingivitis and plaque bacteria vitality. *J Clin Periodontol*. 1990;17:292–7.
- Brex M, MacDonald LL, Legary K, Cheang M, Forgay MGE. Longterm effects of Meridol and chlorhexidine mouthrinses on plaque, gingivitis, staining, and bacterial vitality. *J Dent Res*. 1993;72:1194–7.
- Chilton NW. Conference on equivalence and superiority claims for products for gingivitis and periodontitis. *J Periodontol Res*. 1992;27:305–456.
- Claydon N, Addy M. The use of plaque area and plaque index to measure the effect of fluoride and chlorhexidine toothpastes on 24-h plaque regrowth. *J Clin Periodontol*. 1995;22:540–2.
- Claydon N, Addy M. A 24-h regrowth study to evaluate the plaque inhibitory properties of a proprietary liquid dentifrice. *J Clin Periodontol*. 1999;26:286–8.
- Claydon N, Hunter L, Moran J, Wade W, Kelty E, Mover R, Addy M. A 6-month home-usage trial of 0.1% and 0.2% delmopinol mouthwashes (I). Effects on plaque, gingivitis, supragingival calculus and tooth staining. *J Clin Periodontol*. 1996;23(3 Pt 1):220–8.
- Claydon N, Smith S, Stiller S, Newcombe RG, Addy M. A comparison of the plaque-inhibitory properties of stannous fluoride and low concentration chlorhexidine mouthrinses. *J Clin Periodontol*. 2002;29:1072–7.
- Claydon NC, Addy M, Newcombe R, Moran J. The prevention of plaque re-growth by toothpastes and solutions containing block copolymers with and without polypeptide. *J Clin Periodontol*. 2005;32:545–8.
- Cosyn J, Wyn I, De Rouck T, Collys K, Bottenberg P, Matthijs S, Sabzevar MM. Short-term anti-plaque effect of two chlorhexidine varnishes. *J Clin Periodontol*. 2005;32:899–904.
- Council on Dental Therapeutics. Guidelines for acceptance of chemotherapeutic products for the control of supragingival dental plaque and gingivitis. *J Am Dent Assoc*. 1986;112:529–32.

- Cousido MC, Tomás M, Tomás I, Limeres J, García-Caballero L, Diz P. Effect of a neutralising agent on the evaluation of the antimicrobial activity of chlorhexidine on the bacterial salivary flora. *Arch Oral Biol*. 2008;53:981–4.
- Cousido MC, Tomás Carmona I, García-Caballero L, Limeres J, Alvarez M, Diz P. In vivo substantivity of 0.12% and 0.2% chlorhexidine mouthrinses on salivary bacteria. *Clin Oral Investig*. 2010;14:397–402.
- Daly CG, Highfield JE. Effect of localized experimental gingivitis on early supragingival plaque accumulation. *J Clin Periodontol*. 1996;23:160–4.
- de Albuquerque RF, Head Jr TW, Mian H, Rodrigo A, Müller K, Sanches K, Ito IY. Reduction of salivary *S. aureus* and mutans group streptococci by a preprocedural chlorhexidine rinse and maximal inhibitory dilutions of chlorhexidine and cetylpyridinium. *Quintessence Int*. 2004;35:635–40.
- Deinzer R, Förster P, Fuck L, Herforth A, Stiller-Winkler R, Idel H. Increase of crevicular interleukin 1beta under academic stress at experimental gingivitis sites and at sites of perfect oral hygiene. *J Clin Periodontol*. 1999;26:1–8.
- Deinzer R, Weik U, Kolb-Bachofen V, Herforth A. Comparison of experimental gingivitis with persistent gingivitis: differences in clinical parameters and cytokine concentrations. *J Periodontol Res*. 2007;42:318–24.
- Elworthy A, Greenman J, Doherty FM, Newcombe RG, Addy M. The substantivity of a number of oral hygiene products determined by the duration of effects on salivary bacteria. *J Periodontol*. 1996;67:572–6.
- Fine DH, Furgang D, Barnett ML, Drew C, Steinberg L, Charles CH, Vincent JW. Effect of an essential oil-containing antiseptic mouthrinse on plaque and salivary *Streptococcus mutans* levels. *J Clin Periodontol*. 2000;27:157–61.
- Fine DH, Furgang D, Sinatra K, Charles C, McGuire A, Kumar LD. In vivo antimicrobial effectiveness of an essential oil-containing mouth rinse 12 h after a single use and 14 days' use. *J Clin Periodontol*. 2005;32:335–40.
- Fischman SL. Design of studies to evaluate plaque control agents. *J Dent Res*. 1979;58:2389–95.
- ICH Harmonised Tripartite Guideline. Guideline for good clinical practice E6(R1) Step 4 version dated 10 June 1996. Retrieved at http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1_Guideline.pdf. Accessed on November 2011.
- Gehlen I, Netuschil L, Georg T, Reich E, Berg R, Katsaros C. The influence of a 0.2% chlorhexidine mouthrinse on plaque regrowth in orthodontic patients. *J Orofac Orthop*. 2000;61:138–48.
- Gjerme P, Bonesvoll P, Rolla G. Relationship between plaque-inhibiting effect and retention of chlorhexidine in the human oral cavity. *Arch Oral Biol*. 1974;19:1031–4.
- Grant MM, Creese AJ, Barr G, Ling MR, Scott AE, Matthews JB, Griffiths HR, Cooper HJ, Chapple IL. Proteomic analysis of a noninvasive human model of acute inflammation and its resolution: the twenty-one day gingivitis model. *J Proteome Res*. 2010;9:4732–44.
- Harper PR, Milsom S, Wade W, Addy M, Moran J, Newcombe RG. An approach to efficacy screening of mouthrinses: studies on a group of French products (II). Inhibition of salivary bacteria and plaque in vivo. *J Clin Periodontol*. 1995;22:723–7.
- Harrap GJ. Assessment of the effect of dentifrices on the growth of dental plaque. *J Clin Periodontol*. 1974;1:166–74.
- Herrera D, Roldán S, Santacruz I, Santos S, Masdevall M, Sanz M. Differences in antimicrobial activity of four commercial 0.12% chlorhexidine mouthrinse formulations: an in vitro contact test and salivary bacterial counts study. *J Clin Periodontol*. 2003;30:307–14.
- Herrera D, Santos S, Ferrús J, Barbieri G, Trombelli L, Sanz M. Efficacy of a 0.15% benzydamine hydrochloride and 0.05% cetylpyridinium chloride mouth rinse on 4-day de novo plaque formation. *J Clin Periodontol*. 2005;32:595–603.
- Human D, Fluss SS. The World Medical Association's Declaration of Helsinki: historical and contemporary perspectives. 2001. Available at: http://www.wma.net/en/20activities/10ethics/10helsinki/draft_historical_contemporary_perspectives.pdf. Accessed on 9 Dec 2011.
- Imrey PB, Chilton NW, Pihlstrom BL, et al. Proposed guidelines for American Dental Association acceptance of products for professional, non-surgical treatment of adult periodontitis. Task Force on Design and Analysis in Dental and Oral Research. *J Periodontol Res*. 1994a;29:348–60.
- Imrey PB, Chilton NW, Pihlstrom BL, et al. Recommended revisions to American Dental Association guidelines for acceptance of chemotherapeutic products for gingivitis control. Report of the Task Force on Design and Analysis in Dental and Oral Research to the Council on Therapeutics of the American Dental Association. *J Periodontol Res*. 1994b;29:299–304.
- Jeffcoat M. What is clinical significance? *J Clin Periodontol*. 2002;29:30–2.
- Jenkins S, Addy M, Newcombe R. The effect of triclosan, stannous fluoride and chlorhexidine products on: (II) salivary bacterial counts. *J Clin Periodontol*. 1990;17:698–701.
- Jenkins S, Addy M, Newcombe R. Triclosan and sodium lauryl sulphate mouthwashes (I). Effects on salivary bacterial counts. *J Clin Periodontol*. 1991;18:140–4.
- Jenkins S, Addy M, Wade W, Newcombe RG. The magnitude and duration of the effects of some mouthrinse products on salivary bacterial counts. *J Clin Periodontol*. 1994;21:397–401.
- König J, Storcks V, Kocher T, Bössmann K, Plagmann H-C. Antiplatelet effect of tempered 0.2% chlorhexidine rinse: an in vivo study. *J Clin Periodontol*. 2002;29:207–10.
- Lie MA, Timmerman MF, van der Velden U, van der Weijden GA. Evaluation of 2 methods to assess gingival bleeding in smokers and non-smokers in natural and

- experimental gingivitis. *J Clin Periodontol*. 1998a;25:695–700.
- Lie MA, van der Weijden GA, Timmerman MF, Loos BG, van Steenberghe TJ, van der Velden U. Oral microbiota in smokers and non-smokers in natural and experimentally-induced gingivitis. *J Clin Periodontol*. 1998b;25:677–86.
- Löe H, Theilade E, Jensen SB. Experimental gingivitis in man. *J Periodontol*. 1965;36:177–87.
- Lorenz K, Bruhn G, Heumann C, Hoffmann T, Netuschil L. How to select study designs and parameters to investigate the effect of mouthrinses? Part II: comparisons between the parameters used. *J Physiol Pharmacol*. 2009a;60:85–90.
- Lorenz K, Bruhn G, Netuschil L, Heumann C, Hoffmann T. How to select study designs and parameters to investigate the effect of mouthrinses? Part I: rationale and background. *J Physiol Pharmacol*. 2009b;60:77–83.
- Lorenz K, Mayer D, Bruhn G, Noack B, Brex M, Heumann C, Toutenburg H, Netuschil L, Nagl M, Gottardi W, Hoffmann T. Effect of N-chlorotaurine mouth rinses on plaque regrowth and plaque vitality. *Clin Oral Investig*. 2009c;13:9–14.
- Matheny JL, Abrams H, Johnson DT, Roth GI. Microcirculatory dynamics in experimental human gingivitis. *J Clin Periodontol*. 1993;20:578–83.
- McBain AJ, Bartolo RG, Catrenich CE, Charbonneau D, Ledder RG, Gilbert P. Effects of a chlorhexidine gluconate-containing mouthwash on the vitality and antimicrobial susceptibility of in vitro oral bacterial ecosystems. *Appl Environ Microbiol*. 2003;69:4770–6.
- McClanahan SF, Bartizek RD. Effects of triclosan/copolymer dentifrice on dental plaque and gingivitis in a 3-month randomized controlled clinical trial: influence of baseline gingivitis on observed efficacy. *J Clin Dent*. 2002;13:167–78.
- Moran J, Addy M, Newcombe R. The antibacterial effect of toothpastes on the salivary flora. *J Clin Periodontol*. 1988;15:193–9.
- Moran J, Addy M, Kohut B, Hovliaras CA, Newcombe RG. Efficacy of mouthrinses in inhibiting the development of supragingival plaque over a 4-day period of no oral hygiene. *J Periodontol*. 1994;65:904–7.
- Moran J, Addy M, Wade W, Milson S, McAndrew R, Newcombe RG. The effect of oxidising mouthrinses compared with chlorhexidine on salivary bacterial counts and plaque regrowth. *J Clin Periodontol*. 1995;22:750–5.
- Netuschil L, Reich E, Brex M. Direct measurement of the bactericidal effect of chlorhexidine on human dental plaque. *J Clin Periodontol*. 1989;16:484–8.
- Netuschil L, Weiger R, Preisler R, Brex M. Plaque bacteria counts and vitality during chlorhexidine, meridol and listerine mouthrinses. *Eur J Oral Sci*. 1995;103:355–61.
- Netuschil L, von Ohle C, Brex M. Die vitalfluoreszenztechnik in der plaqueforschung. *Parodontologie*. 1996;7:293–305.
- Netuschil L, Reich E, Unteregger G, Sculean A, Brex M. A pilot study of confocal laser scanning microscopy for the assessment of undisturbed dental plaque vitality and topography. *Arch Oral Biol*. 1998;43:277–85.
- Netuschil L, Hoffmann T, Brex M. How to select the right mouthrinses in periodontal prevention and therapy. Part I. Test systems and clinical investigations. *Int J Dent Hyg*. 2003;1:143–50.
- Newman MG. Design and implementation of clinical trials of antimicrobial drugs and devices used in periodontal disease treatment. *Ann Periodontol*. 1997;2:180–98.
- Nogueira-Filho GR, Toledo S, Cury JA. Effect of 3 dentifrices containing triclosan and various additives. An experimental gingivitis study. *J Clin Periodontol*. 2000;27:494–8.
- Pan P, Barnett ML, Coelho J, Brogdon C, Finnegan MB. Determination of the in situ bactericidal activity of an essential oil mouthrinse using a vital stain method. *J Clin Periodontol*. 2000;27:256–61.
- Pretty IA, Edgar WM, Higham SM. A study to assess the efficacy of a new detergent free, whitening dentifrice in vivo using QLF planimetric analysis. *Br Dent J*. 2004;197:561–6; discussion 551.
- Putt MS, Van der Weijden GA, Kleber CJ, Saxton CA. Validation of a 21-day, partial-mouth gingivitis model for evaluating chemotherapeutic dentifrices. *J Periodontal Res*. 1993;28:301–7.
- Roberts WR, Addy M. Comparison of the in vivo and in vitro antibacterial properties of antiseptic mouthrinses containing chlorhexidine, alexidine, cetyl pyridinium chloride and hexetidine. Relevance to mode of action. *J Clin Periodontol*. 1981;8:295–310.
- Rodrigues SV, Acharya AB, Bhadbhade S, Thakur SL. Hyaluronan-containing mouthwash as an adjunctive plaque-control agent. *Oral Health Prev Dent*. 2010;8:389–94.
- Rundegren J, Hvid EB, Johansson M, Åström M. Effect of 4 days of mouth rinsing with delmopinol or chlorhexidine on the vitality of plaque bacteria. *J Clin Periodontol*. 1992;19:322–5.
- Saxton CA, Van der Ouderaa FJG. The effect of a dentifrice containing zinc citrate and triclosan on developing gingivitis. *J Periodontol*. 1989;24:75–80.
- Schiott CR, Löe H, Jensen SB, Kilian M, Davies RM, Glavind K. The effect of chlorhexidine mouthrinses on the human oral flora. *J Periodontal Res*. 1970;5:84–9.
- Shearer B, Hall P, Clarke P, Marshall G, Kinane DF. Reducing variability and choosing ideal subjects for experimental gingivitis studies. *J Clin Periodontol*. 2005;32:784–8.
- Socransky SS, Haffajee AD. Dental biofilms: difficult therapeutic targets. *Periodontol* 2000. 2002;28:12–55.
- Tatakis DN, Trombelli L. Modulation of clinical expression of plaque-induced gingivitis. I. Background review and rationale. *J Clin Periodontol*. 2004;31:229–38.
- The ADA Council on Scientific Affairs. Guidelines for chemotherapeutic products for control of gingivitis; 2008. Retrieved on http://www.ada.org/sections/scienceAndResearch/pdfs/guide_chemo_ging.pdf.
- Theilade E. The non-specific theory in microbial etiology of inflammatory periodontal diseases. *J Clin Periodontol*. 1986;13:905–11.

- Theilade E, Wright WH, Jensen SB, Løe H. Experimental gingivitis in man. II. A longitudinal clinical and bacteriological investigation. *J Periodontol Res.* 1966;1:1–13.
- Trombelli L, Farina R, Manfrini R, Tatakis DN. Modulation of clinical expression of plaque-induced gingivitis: effect of incisor crown form. *J Dent Res.* 2004a;83:728–31.
- Trombelli L, Scapoli C, Orlandini E, Tosi M, Bottega S, Tatakis DN. Modulation of clinical expression of plaque-induced gingivitis. III. Response of “high responders” and “low responders” to therapy. *J Clin Periodontol.* 2004b;31:253–9.
- Trombelli L, Tatakis DN, Scapoli C, Bottega S, Orlandini E, Tosi M. Modulation of clinical expression of plaque-induced gingivitis. II. Identification of “high-responder” and “low-responder” subjects. *J Clin Periodontol.* 2004c;31:239–52.
- Trombelli L, Scapoli C, Calura G, Tatakis DN. Time as a factor in the identification of subjects with different susceptibility to plaque-induced gingivitis. *J Clin Periodontol.* 2006;33:324–8.
- Trombelli L, Farina R, Minenna L, Carrieri A, Scapoli C, Tatakis DN. Experimental gingivitis: reproducibility of plaque accumulation and gingival inflammation parameters in selected populations during a repeat trial. *J Clin Periodontol.* 2008;35:955–60.
- Tu YK, Blance A, Clerehugh V, Gilthorpe MS. Statistical power for analyses of changes in randomized controlled trials. *J Dent Res.* 2005;84:283–7.
- Tu Y-K, Maddick I, Kellett M, Clerehugh V, Gilthorpe MS. Evaluating the quality of active-control trials in periodontal research. *J Clin Periodontol.* 2006;33:151–6.
- van der Mei HC, White DJ, Atema-Smit J, van de Belt-Gritter E, Busscher HJ. A method to study sustained antimicrobial activity of rinse and dentifrice components on biofilm viability in vivo. *J Clin Periodontol.* 2006;33:14–20.
- van der Weijden GA, Timmerman MF, Danser MM, Nijboer A, Saxton CA, Van der Velden U. Effect of pre-experimental maintenance care duration on the development of gingivitis in a partial mouth experimental gingivitis model. *J Periodontol Res.* 1994;29:168–73.
- von Ohler C, Weiger R, Decker E, Schlagenhauf U, Brex M. The efficacy of a single pocket irrigation on subgingival microbial vitality. *Clin Oral Investig.* 1998;2:84–90.
- Weiger R, Netuschil L, Von Ohle C, Brex M. Microbial vitality of supragingival dental plaque during initial stages of experimental gingivitis in humans. *J Periodontol Res.* 1995;30:204–9.
- Weiger R, Netuschil L, Wester-Ebbinghaus T, Brex M. An approach to differentiate between antibacterial and antiadhesive effects of mouthrinses in vivo. *Arch Oral Biol.* 1998;43:559–65.
- Welk A, Splieth CH, Schmidt-Martens G, Schwahn Ch, Kocher T, Kramer A, Rosin M. The effect of a poly-hexamethylene biguanide mouthrinse compared with a triclosan rinse and a chlorhexidine rinse on bacterial counts and 4-day plaque re-growth. *J Clin Periodontol.* 2005;32:499–505.
- Yates R, Moran J, Addy M, Mullan PJ, Wade WG, Newcombe R. The comparative effect of acidified sodium chlorite and chlorhexidine mouthrinses on plaque regrowth and salivary bacterial counts. *J Clin Periodontol.* 1997;24:603–9.

It is widely recognized that using well-conducted randomized clinical trials (RCTs) to evaluate the efficacy of treatments is beneficial to society in the long run. Harmful or ineffective treatments can be dropped and replaced by effective ones. However, short-term effects of RCTs—the effects before the results are known, essentially ‘side effects’ of RCTs—should also be evaluated (Braunholtz et al. 2001).

- **Placebo effects:** Which arise from the impact of the consent process on patients, might be different for the sorts of patients who tend to agree to participate than for those who tend to refuse (Braunholtz et al. 2001).

When performing and analyzing prospective clinical trials, investigators should recognize the possible implication of a placebo and Hawthorne effects to the results of the trial.

15.1 Terminology

- **Treatment effect:** Altered balance of (main) treatments of differing efficacy (participants vs. others) (Braunholtz et al. 2001).
- **Protocol effect:** Differences in regimens of main treatments (participants vs. others) (Braunholtz et al. 2001).
- **Care effect:** Differences in incidental aspects of treatment or care (participants vs. others). Any ‘care effect’ would be difficult to estimate separately from the protocol effect but would derive from such things as extra follow-up, or extra nursing cover, for data gathering (Braunholtz et al. 2001).
- **Hawthorne effect:** Changes in patient or clinician behaviour (other than as prescribed by the trial protocol) due to increased knowledge or interest or due to feeling ‘observed’. Such changes in behaviour might also differ in extent between participant and non-participant patients of recruiting clinicians (Braunholtz et al. 2001).

15.2 The Placebo Effect

The placebo effect refers to an improvement in outcomes solely from receiving treatment, even if it is an inert drug or sham surgery (Fox et al. 2008).

In clinical practice, when a physician administers a medicine known to be effective in treating patients with a given condition, it is likely that four modes of healing are operative: (a) treatment-induced healing, (b) placebo-induced healing, (c) healing produced by clinician–patient interaction, and (d) self-healing properties of the human organism. The therapeutic response of the patient will be a function of the treatment-induced healing of the medicine, the placebo effect of taking medicine, the therapeutic benefit of receiving clinical attention (including receiving a diagnosis and reassurance), and the disease-fighting powers of the body. Ordinarily, there will be no way to determine, or no point in determining, the respective contributions of these modes of healing for the situation of a particular patient (Miller and Rosenstein 2006).

The second and third modes of healing (i.e. placebo-induced healing and healing produced by clinician–patient interaction) are often lumped together as constituting the placebo effect. The point of doing so is that there is reason to think that both work by similar psychosomatic mechanisms, and it is difficult to tease them apart in clinical practice because treatment is given in the context of a clinician–patient interaction (Miller and Rosenstein 2006).

As to experimental pain, different cortical (prefrontal cortex, anterior cingulate gyrus, insula, supplementary motor area) and subcortical structures (amygdala, periaqueductal grey, thalamus) have been found to be involved in the placebo response, and they seem to differentiate between the sensory and the emotional/affective components of pain signals. PET receptor-binding studies have provided direct evidence that the m-opioid system involving the brain stem and elaborated cortical networks mediates placebo analgesia, thus confirming previous studies on the blockade of placebo analgesia by the opioid antagonist naloxone. It should be noted that other neurochemical systems have been found to contribute to the placebo effect, for example, the dopaminergic system and CCK (Enck et al. 2008).

Compared to the placebo effect, much less is known about the nocebo effect since the induction of a nocebo response represents a stressful and anxiogenic procedure, thus limiting its ethical investigation. The term nocebo ('I shall harm') was introduced in contraposition to the term placebo ('I shall please') by a number of authors in order to distinguish the pleasing from the noxious effects of placebo. If the positive psychosocial context, which is typical of the placebo effect, is reversed, the nocebo effect can be studied. Therefore, it is important to stress that the study of the nocebo effect relates to the negative psychosocial context surrounding the treatment, and its neurobiological investigation is the analysis of the effects of this negative context on the patient's brain and body. As for the placebo effect, the nocebo effect follows the administration of an inert substance, along with the suggestion that the subject will get worse. However, the

term nocebo-related effect can also be used whenever symptom worsening follows negative expectations without the administration of any inert substance (Enck et al. 2008).

15.3 The Hawthorne Effect

When evaluating interventions aimed at improving clinical practice, however, there are a number of non-specific effects that may influence estimations of the effect of an intervention in RCTs. These include positive attention effects, caused by participants knowing that they are the subject of a study, but also negative and demotivating effects, caused by being allocated to a control rather than to an intervention group. These non-specific effects are currently grouped together under the term Hawthorne effect. If these are imbalanced across study groups in a quality improvement trial, the resulting effect estimates may be biased (Verstappen et al. 2004). However, some studies have found that simply monitoring a particular outcome can change study participant behaviour if they know the investigators are interested in that outcome (De Amici et al. 2000).

The Hawthorne effect received its name from a factory called the Hawthorne Works, where a series of experiments were performed between 1924 and 1932 initially studying the effect of lighting on workers' productivity. Because these interventions varied without the intention to increase worker's performance, it was argued that the increase in productivity was the result of personal attention and of the newness of a program (Leonard and Masatu 2006).

It is possible that a physician's diagnostic accuracy can also be subject to the Hawthorne effect. Namely, if a physician were participating in a study, it is possible that he or she could knowingly or unknowingly become more accurate in diagnosing a condition based on the physical examination (Fox et al. 2008). Leonard (2008) argued that the Hawthorne effect represents an exogenous change within individual doctors' quality levels that is uncorrelated with observed or unobserved patient or illness characteristics. The arrival of the research team is an unanticipated

event that induces short-term changes in quality without altering the types of patients and illnesses at a facility (Leonard 2008).

The Hawthorne effect is distinct from an incentive effect in which clinicians alter their behaviour because they suspect they may be penalized or rewarded for their behaviour. Though the literature suggests that the Hawthorne effect may require a ‘perceived demand for performance’, the pure incentive effect can be differentiated from the Hawthorne effect by the fact that, under the Hawthorne effect, individuals eventually return to their pre-observation level of activity even when they remain under observation (Leonard and Masatu 2006).

In a variety of studies, researchers have attributed both long- (Claydon et al. 1996; Gilbert et al. 1998) and short-term (Binney et al. 1996; Owens et al. 1997a, b; Heasman et al. 1998) oral health improvements to the influence of the Hawthorne effect as an unintended consequence of research participation (Feil et al. 2002).

The Hawthorne effect due to participation in a clinical trial is usually more apparent in the first phase of a study but may vary in duration and magnitude but ultimately reduces the ability to detect actual product differences (Binney et al. 1996). What is also important is that the term “Hawthorne effect” has been used frequently to account for gains made by placebo control groups when none were expected. It is cited in the dental and medical literature and has had a greater influence on the placebo control group than the experimental effect had on the experimental group. In a large number of instances, the Hawthorne effect yields statistically significant improvements for the control group over baseline (for review see Feil et al. 2002).

It appears that subjects knowingly participating in oral hygiene studies improve their tooth cleaning irrespective of the product they receive. A cursory study of the data from this present study indicates that a Hawthorne effect occurred with subjects in the control group showing improvement in plaque and gingivitis as the investigation progressed. This phenomenon is a feature of almost all toothbrushing studies and which a potential antiplaque agent must outperform if

any significant treatment effects are to be shown (Owens et al. 1997a, b; Dorfer et al. 2009; Feil et al. 2002; McCracken et al. 2000; Gehlen et al. 2000; Heasman et al. 1998; Claydon et al. 1996; Binney et al. 1996; Aldridge et al. 1995).

Comparatively modest effects on gingivitis of proven antiplaque agents evaluated in placebo-controlled long-term home use studies are not unusual and have been noted with delmopinol (Claydon et al. 1996) and chlorhexidine formulations (Yates et al. 1993). Hawthorne effects will inevitably occur and mask, to some incalculable degree, the true adjunctive effect of the agent. Exceptions appear to have occurred in studies where plaque control was not intended or understood by the participants as the main aim of the study (Chadwick et al. 1991).

The Hawthorne effect is a component of the non-specific effects of trial participation but is not controlled for by usual controlled trial designs. Most clinical trials are unable to quantify the magnitude of the Hawthorne effect because its defining features, such as extra attention by researchers and higher levels of clinical surveillance, apply equally to treatment and control arms. Although the Hawthorne effect should not affect assessment of the difference between intervention and control, it may result in an inflated estimate of effect size in routine clinical settings by overestimating response in both groups (McCarney et al. 2007).

As a consequence, it was proposed the inclusion of a pretest period between screening and the first experimental period to allow a short period for **Hawthorne stabilization** (Heasman et al. 1998).

The Hawthorne effect clearly may have its implications for clinical research and its generalizability to routine practice. If there is a demonstrable benefit from participating in clinical research, *for whatever reason*, then this has implications for good clinical practice and for improving care (McCarney et al. 2007).

In the context of oral health, Feil et al. (2002) suggested that frequency, duration, and effectiveness of home care regimens (all behavioral in nature) might be improved if the Hawthorne effect could be intentionally induced. The Hawthorne

effect was intentionally induced by deceiving subjects into believing they were participating in a toothpaste study. For these experimental subjects, mean plaque scores were 71% at baseline, 54% at 3 months, and 52% at 6 months. On the other hand, the control group plaque scores were 74% at baseline and rose to 78% at 3 months and 79% at 6 months. Since the dentifrice used was an over-the-counter fluoride product, the improvement in plaque scores could only be attributed to the Hawthorne effect—that is, a change in home care behaviour as a consequence of believing one is participating in an experiment. This study demonstrated that, when used intentionally, the Hawthorne effect can alter subjects' oral health behaviour (Feil et al. 2002).

The authors (Feil et al. 2002) recommend that further studies should be conducted to evaluate the usefulness of the Hawthorne effect on other dental populations (restorative and periodontal). For example, a dentist or hygienist could use a 'new' or 'experimental' toothbrush, toothpick, or floss to test its effect on gingival health or implement a 'new' mouthwash for periodontal patients to reduce gram-negative bacteria. Since the Hawthorne effect is efficient to implement for the practitioner, it may be an ideal method to improve oral hygiene for those patients where traditional interventions (oral hygiene instructions and praise) have failed. From a psychological perspective, it would be interesting to evaluate whether after a period of time during which the Hawthorne effect was producing a beneficial outcome, if subjects could be debriefed and encouraged to continue applying the newly learned oral health habits (Feil et al. 2002).

References

- Aldridge JP, Lester V, Watts TL, Collins A, Viberti G, Wilson RF. Single blind studies of the effects of improved periodontal health on metabolic control in type 1 diabetes mellitus. *J Clin Periodontol*. 1995;22:271–5.
- Binney A, Addy M, Owens J, Faulkner J, McKeonrt S, Everatt L. A 3-month home use study comparing the oral hygiene and gingival health benefits of triclosan and conventional fluoride toothpastes. *J Clin Periodontol*. 1996;23:1020–4.
- Braunholtz DA, Edwards SJ, Lilford RJ. Are randomized clinical trials good for us (in the short term)? Evidence for a "trial effect". *J Clin Epidemiol*. 2001;54:217–24.
- Chadwick B, Addy M, Walker DM. The use of hexetidine mouthwash in the management of minor aphthous ulceration and as an adjunct to oral hygiene. *Br Dent J*. 1991;171:83–7.
- Claydon N, Hunter L, Moran J, Wade W, Kelty E, Mover R, Addy M. A 6-month home-usage of 0.1% and 0.2% delmopinol mouthwashes: effects of plaque, gingivitis, supragingival calculus, and tooth staining. *J Clin Periodontol*. 1996;23:220–8.
- De Amici D, Klersy C, Ramajoli F, Brustua L, Politi P. Impact of the Hawthorne effect in a longitudinal clinical study: the case of anesthesia. *Control Clin Trials*. 2000;21:103–14.
- Dorfer CE, Joerss D, Wolff D. A prospective clinical study to evaluate the effect of manual and power toothbrushes on pre-existing gingival recessions. *J Contemp Dent Pract*. 2009;10:1–8.
- Enck P, Benedetti F, Schedlowski M. New insights into the placebo and nocebo responses. *Neuron*. 2008;59:195–206.
- Feil PH, Grauer JS, Gadbury-Amyot CC, Kula K, McCunniff MD. Intentional use of the Hawthorne effect to improve oral hygiene compliance in orthodontic patients. *J Dent Educ*. 2002;66:1129–35.
- Fox NS, Brennan JS, Chasen ST. Clinical estimation of fetal weight and the Hawthorne effect. *Eur J Obstet Gynecol Reprod Biol*. 2008;141:111–4.
- Gehlen I, Netuschil L, Berg R, Reich E, Katsaros C. The influence of a 0.2% chlorhexidine mouthrinse on plaque regrowth in orthodontic patients. A randomized prospective study. Part I: clinical parameters. *J Orofac Orthop*. 2000;61:54–62.
- Gilbert GH, Duncan RP, Campbell AM. Evaluation for an observation effect in a prospective cohort study of oral health outcomes. *Community Dent Oral Epidemiol*. 1998;26:233–40.
- Heasman P, Wilson Z, Macgregor I, Kelly P. Comparative study of electric and manual toothbrushes in patients with fixed orthodontic appliances. *Am J Orthod Dentofacial Orthop*. 1998;114:45–9.
- Leonard KL. Is patient satisfaction sensitive to changes in the quality of care? An exploitation of the Hawthorne effect. *J Health Econ*. 2008;27:444–59.
- Leonard K, Masatu MC. Outpatient process quality evaluation and the Hawthorne effect. *Soc Sci Med*. 2006;63:2330–40.
- McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne effect: a randomised, controlled trial. *BMC Med Res Methodol*. 2007;7:30.
- McCracken GI, Heasman L, Stacey F, Kelly PJ, Heasman PA. Testing the efficacy of plaque removal of a prototype brush head for a powered toothbrush. *J Clin Periodontol*. 2000;27:542–8.
- Miller FG, Rosenstein DL. The nature and power of the placebo effect. *J Clin Epidemiol*. 2006;59:331–5.

- Owens J, Addy M, Faulkner J, Lockwood C, Adair R. A short-term clinical study design to investigate the chemical plaque inhibitory properties of mouthrinses when used as adjuncts to toothpastes: applied to chlorhexidine. *J Clin Periodontol*. 1997a;24:732–7.
- Owens J, Addy M, Faulkner J. An 18-week home-use study comparing the oral hygiene and gingival health benefits of triclosan and fluoride toothpastes. *J Clin Periodontol*. 1997b;24:626–31.
- Verstappen WH, van der Weijden T, ter Riet G, Grimshaw J, Winkens R, Grol RP. Block design allowed for control of the Hawthorne effect in a randomized controlled trial of test ordering. *J Clin Epidemiol*. 2004;57:1119–23.
- Yates R, Jenkins S, Newcombe R, Wade W, Moran J, Addy M. A 6-month home usage trial of a 1% chlorhexidine toothpaste (1). Effects on plaque, gingivitis, calculus and toothstaining. *J Clin Periodontol*. 1993;20:130–8.

Halitosis is a general term used to define an unpleasant or offensive odour emanating from the breath regardless of whether the odour originates from oral or non-oral sources. Other terms used are bad or foul breath, breath malodour, oral malodour, foetor ex ore, and foetor oris (van den Broek et al. 2007). Halitosis is widespread and is believed to affect one-quarter of the population around the world, and most people have this condition from time to time. Those affected are usually unaware of their condition possibly because of smell adaptation. Another explanation could be the pathways between the inhaled and exhaled air diverge—because the expelled air from the mouth travels horizontally, whereas the air breathed in travels primarily vertically; there is a lowered chance of detecting the smell from the expelled air (Lee et al. 2004, 2007).

A useful classification system for different types of halitosis was published by Yaegaki and Coil (2000a). This classification system comprises three categories: pseudo-halitosis, halitophobia, and genuine halitosis. **Genuine halitosis** was subclassified as physiologic halitosis and pathologic halitosis. Pathologic halitosis was further categorized to oral pathologic halitosis and extra-oral pathologic halitosis. If oral malodour does not exist but the patient believes that he or she has oral malodour, the diagnosis would be **pseudo-halitosis**. If, after

treatment for either genuine halitosis or pseudo-halitosis, the patient still believes that he or she has halitosis, the diagnosis would be halitophobia. This classification allows the clinician to diagnose a psychological condition. Pseudo-halitosis cannot be treated by dental practitioners, and halitophobia patients must be referred to psychological specialists (Yaegaki and Coil 2000a).

16.1 Etiopathogenesis of Oral Malodour

16.1.1 Halitosis from Oral Causes

In about 85% of patients with persistent genuine halitosis, the odour originates from the mouth, mainly from microorganisms (Scully and Greenman 2008). Oral malodour arises from the production of volatile malodourous compounds by the action of bacteria in breaking down components of epithelial cells, salivary and serum proteins, and food debris. A wide range of molecular species are able to contribute to the overall production of malodour. Notably, much attention has been given to volatile sulphur compounds, including methyl mercaptan, hydrogen sulphide, and dimethyl sulphide. However, a number of other

compounds may also contribute to malodour (Table 16.1) (Hughes and McNab 2008).

The production of these compounds is mediated by the putrefaction of debris and protein substrates by a wide range of oral anaerobic organisms, particularly those which exhibit proteolytic activity. It is likely that there is a complex interaction between several oral bacteria species (mainly gram-negative anaerobic flora) because no single specific bacterial infection has invariably been associated with halitosis (Scully and Greenman 2008; Haraszthy et al. 2007). In the case of VSC generation, bacteria act on the sulphur-containing amino acids cysteine, cystine, and methionine which are available following proteolysis (Hughes and McNab 2008 and references therein).

Persistent halitosis occurring as a result of intra-oral causes usually originates from the posterior dorsum of the tongue and/or oral/dental diseases, including periodontal disease, and can

be severe enough to be considered socially unacceptable (also known as pathologic halitosis). It is most likely to occur in patients with conditions that favour the accumulation of food and bacterial plaque on intra-oral surfaces (teeth, gingiva, and mucosal tissues, especially the dorsum of the tongue) and the development of anaerobic ecosystems. Predisposing factors include poor oral hygiene, gingival and periodontal disease, disorders of the oral mucosa, reduced salivary flow, and the wearing of dental appliances (Table 16.2) (Scully and Greenman 2008).

The role of periodontitis in the aetiology of oral malodour is supported by a number of studies reviewed by Morita and Wang (2001a). The proportion of periodontal pathogens, such as *Porphyromonas gingivalis*, *Prevotella intermedia*, *Tannerella forsythia*, and *Treponema denticola* that are able to produce high levels of H₂S and CH₃SH, detected on the tongue has been reported to be significantly correlated with VSC levels or periodontal pocket depth (Tanaka et al. 2004). It has been also demonstrated that saliva from patients with periodontitis may produce increased amounts of VSC (Yaegaki and Sanada 1992a; Awano et al. 2002), that VSC levels are elevated in deep periodontal pockets compared with shallow periodontal pockets (Solis-Gaffar et al. 1980; Coli and Tonzetich 1992; Morita and Wang 2001b, c), and that VSCs in mouth air increase with the severity of periodontal disease (Tonzetich 1978; Yaegaki and Sanada 1992a; Yaegaki and Sanada 1992b; Tsai et al. 2008; Takeuchi et al. 2010). Organoleptic testing

Table 16.1 Volatile molecules contributing to oral malodour (Hughes and McNab 2008) (Reprinted with permission from Elsevier)

Volatile sulphur compounds (VSCs)	
	Methyl mercaptan, hydrogen sulphide and dimethyl sulphide
Diamines	
	Putrescine, cadaverine
	Short-chain fatty acids
	Butyric acid, propionic acid
Phenyl compounds	
	Indole, skatole, pyridine

Table 16.2 Main oral causes of halitosis (Scully and Greenman 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Plaque-related gingival and periodontal disease	Gingivitis, periodontitis, acute necrotizing ulcerative gingivitis, pericoronitis, abscesses
Ulceration	Systemic disease (inflammatory/infectious disorders, cutaneous, gastrointestinal and hematological disease), malignancy, local causes, aphthae, drugs
Hyposalivation	(e.g. from drugs, Sjögren's syndrome, radiotherapy, chemotherapy)
Tongue coating	Poor hygiene
Wearing dental appliances	Poor hygiene
Dental conditions	Food packing
Bone diseases	Jaw dry sockets, osteomyelitis, osteonecrosis, malignancy

significantly correlated with periodontal probing depth and a percentage of periodontal pocket depth ≥ 4 mm ($r=0.40$ and 0.39 , respectively). There were significant correlations between methyl mercaptan/hydrogen sulphide ratio and periodontal parameters (Takeuchi et al. 2010). Furthermore, for subjects with periodontitis, there were statistically significant reductions in oral malodour after periodontitis treatment or tongue cleaning; however, major reductions were found after periodontitis treatment. For those in the gingivitis group, there were also statistically significant reductions in oral malodour after gingivitis treatment or tongue cleaning, but the most marked reductions were observed after tongue cleaning (Pham et al. 2011). These studies suggest that periodontal disease is one of the causes of oral malodour.

As it has indicated that patients with periodontal disease have increased tongue coating than healthy controls (Yaegaki and Sanada 1992a, b), it was also suggested that the reported association between oral malodour and periodontitis is primarily due to the effects of periodontal disease on tongue coating (Hughes and McNab 2008). In support of this, a number of studies have not found evidence of strong direct association between periodontitis and halitosis (Bosy et al. 1994; Stamou et al. 2005; John and Vandana 2006). In these studies, however, portable devices were employed for measuring VSCs, and a small number of subjects were enrolled, which might affect the evaluation of the association between periodontal disease and oral malodour (Takeuchi et al. 2010).

The potential importance of volatile sulphur compounds (VSC) in the transition of periodontal tissues from clinical health to gingivitis and then to periodontitis was also emphasized (Ratcliff and Johnson 1999). VSC may play a role in the pathogenesis of inflammatory conditions such as periodontitis. Since these compounds result from bacterial putrefaction of protein, investigations have been conducted to determine whether specific bacteria are associated with odour production. Two members of this family, hydrogen sulphide (H_2S) and methyl mercaptan (CH_3SH), are primarily responsible for mouth odour. Direct exposure to either of these

metabolites adversely affects protein synthesis by human gingival fibroblasts in culture. However, methyl mercaptan has the greatest effect. Other *in vitro* experiments have demonstrated that cells exposed to methyl mercaptan synthesize less collagen, degrade more collagen, and accumulate collagen precursors which are poorly cross-linked and susceptible to proteolysis. CH_3SH also increases permeability of intact mucosa and stimulates production of cytokines which have been associated with periodontal disease. VSCs, and in particular methyl mercaptan, are therefore capable of inducing deleterious changes in both the extracellular matrix and the local immune response of periodontal tissues to plaque antigens (Ratcliff and Johnson 1999). Furthermore, it has been demonstrated that H_2S inhibited the proliferation of human osteoblastic cells through the MAPK pathway (Imai et al. 2009), induces osteoclast differentiation without receptor activator of nuclear factor κB ligand (Li et al. 2010), and induces apoptosis in epithelial cells derived from human gingiva (Murata et al. 2008).

16.1.2 Halitosis from Extra-oral Causes

Halitosis is less frequently associated with extra-oral causes (i.e. conditions and diseases that do not affect primarily the oral cavity). Respiratory disorders (of the nose, sinuses, tonsils, and pharyngeal regions), as well as diseases of the gastrointestinal system, can result in the presence of odiferous gases in the air expelled from the oral cavity and the nose (Table 16.3) (Scully and Greenman 2008).

16.2 Assessment Methods of Oral Malodour

The three primary measurement methods of genuine halitosis are organoleptic measurement, gas chromatography, and sulphide monitoring. Additional or alternative measurement methods are BANA test, chemical sensors, salivary incubation

Table 16.3 Extra-oral causes of halitosis (Scully and Greenman 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Respiratory system (microbial etiology)	Sinusitis
	Antral malignancy
	Cleft palate
	Foreign bodies in the nose
	Nasal malignancy
	Tonsilloliths
	Tonsillitis
	Pharyngeal malignancy
	Lung infections
	Bronchitis
	Bronchiectasis
	Lung malignancy
Gastrointestinal tract	Esophageal diverticulum
	Gastro-esophageal reflux disease
	Malignancy
Metabolic disorders (blood borne)	Acetone-like smell in uncontrolled diabetes
	Uremic breath in renal failure
	Foetor hepaticus in liver disease
	Trimethylaminuria (fish odor syndrome)
	Hypermethioninemia
Drugs (blood borne)	Cystinosis
	• Amphetamines
	• Chloral hydrate
	• Cytotoxic agents
	• Dimethyl sulphoxide
	• Disulfiram
	• Nitrates and nitrites
Psychogenic causes	• Phenothiazines
	• Solvent abuse

test, quantifying β -galactosidase activity, ammonia monitoring, ninhydrin method, and polymerase chain reaction (van den Broek et al. 2007).

Organoleptic or hedonic measurement is a simple commonly used measurement method of halitosis by an examiner. A plastic tube is inserted into the patient's mouth, preventing the dilution of mouth air with room air. While the patient is exhaling slowly, the examiner judges the odour at the other end of the tube. A privacy screen with a hole for the straw or the tube can be used to separate the examiner from the patient. Nasal-breath

odour can be measured with a tube inserted into one of the nostrils, while the other nostril is closed by a finger. Various scoring systems can be used for estimating the intensity of the odour. The most widely used scale is ranging from 0 to 5: 0=no odour, 1=barely noticeable odour, 2=slight but clearly noticeable odour, 3=moderate odour, 4=strong odour, 5=extremely foul odour (van den Broek et al. 2007).

Gas chromatography, performed with apparatus equipped with a flame photometric detector, is specific for detecting sulphur in mouth air. Gas chromatography is considered the gold standard for measuring oral malodour because it is specific for volatile sulphur compounds, the main cause of oral malodour (Yaegaki and Coil 2000b).

Sulphide monitors analyze for total sulphur content of the subject's mouth air. Although compact sulphide monitors are portable and easy to use, most are not specific for volatile sulphur compounds. For example, the Halimeter (Interscan, Chatsworth, CA) has high sensitivity for hydrogen sulphide, but low sensitivity for methyl mercaptan, which is a significant contributor to halitosis caused by periodontal disease (Yaegaki and Coil 2000b).

Measurement of halitosis is complicated by a variety of parameters, and each method has specific advantages and shortcomings with respect to these parameters. Organoleptic measurement by a panel of sensory judges and gas chromatography is very reliable, but expensive and not very practical methods. Nevertheless, the use of organoleptic measurement is suggested as the 'gold standard' or primary indicator of halitosis. This statement is corroborated by research findings with the comment that because of the inherent subjectivity, it should not be the sole measurement method in defining patients with halitosis 118. Gas chromatography is the preferable method if precise measurements of specific gases are required, but the method cannot be easily clinically implemented since it requires relatively high cost, highly trained persons, and extensive procedures. Sulphide monitoring is a relatively inexpensive and easily used method but has the limitation that important odours are not detected. Further improvement of the three

primary measurement methods of halitosis should be one of the aims of halitosis research (van den Broek et al. 2007).

The Diamond Probe®/Perio 2000® System (Diamond General Development Corp., www.lifesciencesworld.com) reportedly detects periodontal disease during routine dental examinations by measuring relative sulphide concentrations as an indicator of gram-negative bacterial activity. The system consists of a single-use disposable probe tip with micro-sensors connected to a main control unit. The probe might detect periodontal disease at an early stage and might find an active site that requires treatment. However, the probing pressure is not controlled. Also, periodontal disease can be caused by bacteria that do not produce volatile sulphur compounds, creating the potential for some disease activity to be missed (Ramachandra et al. 2011).

16.3 Treatment of Oral Malodour

The management includes first determining which cases may have an extra-oral aetiology. A full oral examination is indicated, and if an oral cause is likely or possible, management should include treatment of the cause, and other measures (Scully and Felix 2005).

In cases of malodour which may have an extra-oral aetiology, the responsibility of the general dental practitioner is to refer the patient for evaluation to a specialist. This may involve an oral medicine opinion, an otorhinolaryngologist to rule out the presence of chronic tonsillitis or chronic sinusitis, a physician to rule out gastric, hepatic, endocrine, pulmonary, metabolic, or renal disease, or a psychologist or psychiatrist (Scully and Felix 2005; Delanghe et al. 1997).

The available methods leading to lowering of oral malodour level can be divided into the following: usage of masking products, mechanical reduction of microorganisms and their substrates, chemical reduction of microorganisms, and chemical neutralization of odorous compounds, including volatile sulphur-containing compounds. Patients diagnosed as suffering from non-oral halitosis should be referred to a clinic for

otorhinolaryngology or internal medicine for appropriate treatment (van den Broek et al. 2008; Quirynen et al. 2003).

Masking products are not effective in reducing microorganisms or their substrates and in neutralizing odorous compounds. With some masking products, such as menthol and mint, only a short-term masking effect of <2 h of halitosis can be expected. Chemical reduction of microorganisms by antimicrobial ingredients in oral health-care products is only temporary effective. The effectiveness of the ingredients is dependent on their concentration, and above a certain concentration, the ingredients may have unpleasant side effects. Good short-term results were reported with chlorhexidine. Triclosan seems less effective. Essential oils and cetylpyridinium chloride are only effective for short-time periods of up to 2 or 3 h. Allylpyrocatechol, L-trifluoromethionine, and dehydroascorbic acid could be promising antibacterial agents. No clinical trials were found comparing the effect of antimicrobial ingredients in oral health-care products with the effect of mechanically reducing bacteria and substrates (van den Broek et al. 2008). Chemical ingredients of oral health-care products seem most effective when applied in addition to instructions in oral hygiene and professional mechanical cleaning (Quirynen et al. 2005; Roldán et al. 2005; van den Broek et al. 2008).

Metal ions and oxidizing agents are active ingredients of oral health-care products for chemically neutralizing volatile sulphur-containing compounds. Again, the effectiveness of the active ingredients is dependent on their concentration, and above a certain concentration, the ingredients can have unpleasant side effects. Zinc seems to be an effective safe metal at concentrations of at least 1%. Oxidizing agents, such as hydrogen peroxide, ClO_2 , and iminium are reported to be effective for longer periods of time in patients with slightly unpleasant halitosis. These agents could be used in addition to daily mechanical cleaning in order to reach a day-long effect (van den Broek et al. 2008).

Relevant enhanced or synergistic effects of combinations of chemical agents and ingredients in oral health-care products were demonstrated for:

- Chlorhexidine–zinc
- Chlorhexidine–cetylpyridinium chloride–zinc
- Sodium–zinc (exclusively when a fresh mixture is used)
- Iminium–zinc (van den Broek et al. 2008)

Table 16.4 shows a comparison of various oral malodour therapies or treatments made on the basis of objective reduction of volatile sulphur compound or component gases (by gas chromatography, Halimeter, or sensor) in comparison with appropriate controls (trials of >10 subjects) (Scully and Greenman 2008).

Tongue brushing on a regular basis, particularly aiming at removing the coating on the dorsum of the tongue, has been found to be fruitful in reducing oral malodour (range 59–88%). Tongue cleaning can be carried out by using a modern tongue scraping instrument that is available and that consists of a long strip of plastic ribbon, which is held in both hands and bent so that the edge can be pulled down over the dorsal surface of the tongue removing the coating. It has also been mentioned that the inverted bowl of a spoon may be used as a substitute for the commercial variety of tongue scrapers. Brushing also appears to be an easy method of cleaning the tongue, provided that the gagging reflex can be controlled (Danser et al. 2003). Suggestions for tongue cleaning procedure were presented (Christensen 1998).

The Cochrane systematic review promotes evidence-based outcome studies. A recent review was conducted by Outhouse et al. (2006) to determine reliable evidence concerning the effectiveness of tongue scraping or cleaning, compared with other interventions for controlling halitosis. The review included two trials involving a total of 40 participants. Based on the independent data from these two trials, the tongue cleaner or the tongue scraper demonstrated a statistically significant difference in reducing levels of volatile sulphur compounds (VSCs) when compared with the toothbrush. The findings indicate a small but statistically significant difference in reduction of VSC levels when tongue scrapers or cleaners, rather than toothbrushes, are used to reduce halitosis in adults (Outhouse et al. 2006). Similarly, Van der Sleen et al. (2010) demonstrated that mechanical approaches, such as tongue brushing or tongue scraping to clean the dorsum

of the tongue, have the potential to successfully reduce breath odour and tongue coating.

16.4 Guidelines for Investigating Subjects with Halitosis

Many publications in the area of halitosis research do not clearly define the type of halitosis under investigation, and no consensus on standards for screening and assessment exists (Donaldson et al. 2007). Differing protocols have been used, without uniform agreement, for examining subjects with halitosis (Schmidt and Tarbet 1978; Rosenberg et al. 1991; Yaegaki and Coil 2000b; Kazor et al. 2003; Roldan et al. 2003).

Yaegaki and Coil (2000b) published guidelines for investigating subjects with halitosis. The recommended examination procedures are described below. Patients were instructed to abstain from taking antibiotics for 3 weeks before the assessment, to abstain from eating garlic, onion, and spicy foods for 48 h before the assessment, and to avoid using scented cosmetics for 24 h before the assessment. Patients were instructed to abstain from ingesting any food or drink, to omit their usual oral hygiene practices, to abstain from using oral rinse and breath fresheners, and to abstain from smoking for 12 h before the assessment. The oral malodour examiner, who should have a normal sense of smell, was required to refrain from drinking coffee, tea, or juice and to refrain from smoking and using scented cosmetics before the assessment (Yaegaki and Coil 2000b). However, in their published guidelines for investigating subjects with halitosis Yaegaki and Coil (2000b) did not specify the type of halitosis under investigation and therefore no exclusion criteria were documented (Donaldson et al. 2007).

A detailed clinical protocol for screening and assessing subjects with a complaint of halitosis, in order to recruit patients with halitosis arising from within the oral cavity, which is not related to generalized chronic gingivitis, chronic periodontitis, or pathology of the mucous membranes, was described by Donaldson et al. (2007). An extensive list of exclusion criteria was applied at the initial visit: poor oral hygiene (generalized visible plaque and calculus deposits), generalized chronic gingivitis (visual

Table 16.4 Comparison of various oral malodor therapies or treatments made on the basis of reduction of volatile sulphur compounds (VSC) or component gases (by gas chromatography, halimeter or sensor) in comparison with appropriate controls (trials of $n > 10$ subjects) (Scully and Greenman 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Category of treatments Clinical therapy alone	Volunteers per group	Comments	Outcomes	Instrument used to measure VSC or gas	Statistical significance of reduction	Reference
POHC and mouthwash (clinic)	$n = 923$	Post-treatment assessments following visits to bad breath clinic	98% of population showed reductions of VSC from >180 p.p.b. to <180	Halimetry	NA	Richter (1996)
POHC	$n = 37$	Two-year study; test group received POHC every week including tongue cleaning. Control group ($n = 30$) did not	60% reduction in CH_3SH when measured using an electronic sensor sensitive to CH_3SH	Electronic	$P < 0.05$	Adachi et al. (2002)
POHC (clinic)	$n = 92$	Before vs. after clinical treatment at bad breath clinic	80% reduction in VSC	Gas chromatography	$P < 0.0001$	Tanaka et al. (2003)
Short-term (0–8 h) effects (single use)						
Tongue cleaning	$n = 30$	Three-hour, randomized crossover study (three groups)	42% reduction in VSC immediately after treatment and still significantly reduced for up to 25 min thereafter	Halimetry	$P < 0.001$	Seemann et al. (2001)
Zn^{2+} mouthrinse	$n = 62$	Three-hour, parallel-study (three groups), Zn^{2+} vs. controls	70–80% reduction of VSC at 1, 2 and 3 h	Gas chromatography	$P < 0.05$	Schmidt et al. (1978)
Zn^{2+} (toothpaste)	$n = 11$	Three-hour, crossover study; five treatments (+ 1-week washout periods) (silica base dentrifice with 2% Zn)	Between 40 and 70% reduction for both H_2S and CH_3SH in the Zn pastes compared with controls at 1, 2 and 3 h post-treatment	Gas chromatography	$P < 0.05$	Brunette et al. (1998)
ClO_2 (mouthwash)	$n = 16$	96-h parallel study with two groups: test (ClO_2 at 0.1%) and control (water; $n = 15$)	21% and 19% reduction in VSC at 2 h and 4 h respectively	Halimetry	$P < 0.01$	Frascella et al. (2000)
Two-phase rinse (essential oils plus aqueous CPC), Chlorhexidine gluconate (0.2%)	$n = 19$	One-day, parallel study with three groups (two-phase mouthwash, CHX and placebo)	For two-phase rinse, 38% reduction in VSC at 8–10 h compared with placebo. For CHX, 50% reduction in VSC at 8–10 h compared with placebo	Halimetry	$P < 0.05$ $P < 0.05$	Rosenberg et al. (1992)

(continued)

Table 16.4 (continued)

Category of treatments	Volunteers per group	Comments	Outcomes	Instrument used to measure VSC or gas	Statistical significance of reduction	Reference
Mouthrinses: CHX (0.4%), CHX (0.05%), Triclosan, (0.03%), CPC (0.05%)	n=12	A six-step double-blind, crossover, randomized study on morning breath in volunteers with healthy periodontium (who refrained from mechanical plaque control during 4-day periods).	For CHX (0.12%), a 63% reduction in VSC. For CHX (0.2%), a 70% reduction in VSC. For other agents, reductions in VSC were not statistically significant.	Halimetry	P=0.01 P=0.001	Carvalho et al. (2004)
Long-term or cumulative effect (regular use)						
Sn ²⁺ (toothpaste) (5-day effect)	n=96	Five-day, randomized, controlled, examiner blind, parallel study of four groups (n=96 in each group):	50–60% reduction for VSC for Sn-containing paste compared with control (reduction increases after cumulative use)	Halimetry	P<0.05	Gerlach et al. (1998)
CHX + CPC + Zn ²⁺ Mouthrinse (14-day effect)	n=40	Two-week double-blind parallel study of test product (CHX + CPC + Zn) vs. placebo control. Reduction measured on day 14	45% reduction in VSC	Halimetry	P<0.01	Winkel et al. (2003)

Exclusions: trials that have fewer than 11 subjects; those that only use organoleptic or hedonic methods; and those where subjects have known inflammatory periodontal disease
CH₃SH methyl mercaptan, CHX chlorhexidine, ClO₂ chlorine dioxide, CPC cetylpyridinium chloride, GC gas chromatography, H₂S hydrogen sulphide, POHC professional oral healthcare, Sn stannous, VSC volatile sulphur compounds, Zn zinc

signs of gingivitis), generalized chronic periodontitis (clinical probing depths of ≥ 5 mm; more than five sites with clinical probing depths of >3 mm but ≤ 5 mm), and caries (cavitation in one or more teeth that may cause food trapping); pathology of the oral mucous membranes or attached gingivae; diseases of the respiratory tract, including sinus disorders and asthma; diabetes mellitus, kidney, liver, or stomach disorders, and HIV/AIDS; Sjögren's syndrome; antibiotic therapy in the previous 4 weeks; prescribed medication that can cause xerostomia as listed in the British National Formulary; and edentulousness and smoking (Donaldson et al. 2007).

Eligible subjects were asked to follow strict instructions and complete a questionnaire prior to their second visit for data collection. The purpose of the questionnaire was to verify the general health status of the participants and attempt to exclude any psychosomatic cause of halitosis (Donaldson et al. 2007).

Participant instructions: For 48 h prior to the halitosis assessment, subjects were asked to avoid the following: (1) eating foods containing garlic, onions, and strong spices; (2) consuming alcohol; and (3) using mouthwashes. On the morning of the appointment, subjects were asked to refrain from drinking coffee, eating mints, using minted chewing gum, or scented oral hygiene products and to avoid wearing heavily scented perfumes or aftershaves. This was essential in preventing dietary or cosmetic odours from influencing organoleptic and Halimeter assessments. Subjects were asked to have a light breakfast a minimum of 2 h prior to the breath odour assessment and to brush with water to remove overnight plaque deposits and food debris prior to the examination (Donaldson et al. 2007).

Examiner instructions: The examiner followed the same pre-halitosis assessment instructions as the participants for 24 h prior to undertaking the examinations, the only exception being the use of fluoridated toothpaste (Donaldson et al. 2007).

16.5 The ADA Guidelines on Oral Malodour Products

For over 130 years, the American Dental Association (ADA) has been an important information source on the safety and effectiveness of dental products

both for the consumer and the clinician. The ADA Seal of Acceptance Program has assisted the public and profession in making decisions regarding dental therapeutics, materials, instruments, and equipment by its systematic review of product evaluations and testing (Wozniak 2005).

Late in 2003, the council completed the development of Acceptance Program Guidelines for Products Used in the Management of Oral Malodor (http://www.ada.org/sections/scienceAndResearch/pdfs/guide_oral_malodor.pdf). The guidelines as published apply to products that are designed to manage oral malodour of non-systemic origin. This type of oral malodour is generated by microorganisms or metabolic compounds that reside on the teeth, tongue, or other areas in the oral cavity. Such malodour corresponds to approximately 90% of the observed cases. Products that manage such malodour by either chemical agents or mechanical means may be considered under these guidelines (Wozniak 2005).

16.5.1 Safety of Oral Malodour Products

Regarding safety, the guidelines require that a 6-month study should be conducted unless the product has already been used for plaque and gingivitis control or whose active ingredient is generally recognized as safe. If not, oral flora should be monitored over a 6-month period in appropriately sized clinical study to determine if development of opportunistic and pathogenic organisms occurs. In addition, effects on oral soft and hard tissues should be assessed (Wozniak 2005):

(A) Effect on oral soft tissues

- (i) Gingivitis. Since some chemical agents may cause an increase in pathogenic organisms, gingival inflammation should be assessed with an appropriate index, for example, Loe and Silness.
- (ii) Oral soft tissues. Evidence should be provided that the product does not adversely affect oral soft tissues, including staining, allergic reaction, oral ulceration, candidiasis, or secondary infections of the oral mucosa that may be manifestations of the proliferation of opportunistic microorganisms.

(B) Effect on hard tissues and restorative materials

Evidence of lack of effect of products on hard tissues and restorative materials (staining, shade alterations, and loss of structure) should be provided.

(C) Toxicology

Information submitted for products containing chemical agents shall include assessments of possible toxic effects of the active agent or adverse effects of the product formulation.

(D) Patient-perceived adverse effects

Data should be provided on the effect of the product, if any, regarding patient reports of changes in taste, changes in saliva flow, burning sensation, xerostomia, or other characteristics that may be unique to the product.

(E) Microbiology

Oral flora should be monitored in subjects during the study for the development of opportunistic and pathogenic organisms. Data should be obtained at baseline, 3 weeks and 6 months.

16.5.2 Efficacy of Oral Malodour Products

According to the ADA guidelines, to demonstrate efficacy in the management of oral malodour, two independent 3-week clinical studies utilizing an appropriate placebo control should be conducted. Either crossover or parallel group designs are permitted. Measurements of oral malodour should be performed at a minimum of two appropriate time periods after baseline during the 3-week period. Significant reductions in oral malodour from baseline to the subsequent time points relative to the placebo control should be demonstrated. In addition, 80% of the subjects should demonstrate a reduction to questionable or no oral malodour at some time during the treatment period. Furthermore, evidence should be provided from the previously discussed 6-month clinical study that development of microbial resistance does not occur. Also, mechanism of action should be given (if known) (Wozniak 2005).

16.5.3 Clinical Protocol Guidelines Regarding Clinical Trials for Products Used in the Management of Oral Malodour

The guidelines also give substantial information regarding the design of the clinical trials that should be conducted to demonstrate efficacy (Wozniak 2005).

- Regarding protocol: For active chemical agents (type 1 products), in each study, the active product should be compared with an appropriate placebo control, along with regular oral hygiene (brushing and flossing). For mechanical means (type 2 products), the control should be regular oral hygiene.
- Regarding subject selection: Individuals should be included who have intrinsic malodour of oral origin and have an average organoleptic intensity rating of at least 2.0 on a 0–5 intensity scale. Subjects with oral diseases such as advanced periodontitis and subjects who smoke or wear oral appliances should be excluded.

Regarding study design: Four hours prior to evaluation of oral malodour, the subjects must abstain from ingestion of food, drinking, and oral hygiene.

- Regarding study duration: The efficacy studies will be conducted for a minimum of 3 weeks using a crossover or parallel design utilizing at least one appropriate control. Safety and microbiological assessments should be continued for 6 months from one clinical effectiveness trial if appropriate. Products should continue to be used (depending on product instructions) for at least 3 weeks and demonstrate continued effectiveness over that time period.
- Regarding assessments:
 - Safety assessments: (1) microbiological assessment; (2) oral soft tissue assessment (gingival inflammation should be assessed at baseline, 3 weeks, and at 6 months with an appropriate index, e.g. Loe and Silness); (3) oral hard tissue assessment.
 - Efficacy assessments: (1) Assessment of malodour: Organoleptic intensity or hedonic examinations will be performed by at least two

properly trained and calibrated odour judges blinded to each other, based on samples obtained from oral air. Two types of instruments that may be used for detection of volatile compounds are gas chromatography and portable sulphide monitors. (2) Microbiological assessment: Evidence should be provided from a 6-month clinical study that this does not occur with use of this product. Data should be obtained at baseline, 3 weeks, and again at 6 months.

References

- Adachi M, Ishihara K, Abe S, Okuda K, Ishikawa T. Effect of professional oral health care on the elderly living in nursing homes. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2002;94:191–5.
- Awano S, Gohara K, Kurihara E, Ansai T, Takehara T. The relationship between the presence of periodontopathogenic bacteria in saliva and halitosis. *Int Dent J.* 2002;52 Suppl 3:212–6.
- Bosy A, Kulkarni GV, Rosenberg M, McCulloch CA. Relationship of oral malodor to periodontitis: evidence of independence in discrete subpopulations. *J Periodontol.* 1994;65:37–46.
- Brunette DM, Proskin HM, Nelson BJ. The effects of dentifrice systems on oral malodor. *J Clin Dent.* 1998;9:76–82.
- Carvalho MD, Tabchoury CM, Cury JA, Toledo S, Nogueira-Filho GR. Impact of mouthrinses on morning bad breath in healthy subjects. *J Clin Periodontol.* 2004;31:85–90.
- Christensen GJ. Why clean your tongue? *J Am Dent Assoc.* 1998;129:1605–7.
- Coli JM, Tonzetich J. Characterization of volatile sulphur compounds production at individual gingival crevicular sites in humans. *J Clin Dent.* 1992;3:97–103.
- Danser MM, Gómez SM, Van der Weijden GA. Tongue coating and tongue brushing: a literature review. *Int J Dent Hyg.* 2003;1(3):151–8.
- Delanghe G, Ghyselen J, van Steenberghe D, Feenstra L. Multidisciplinary breath-odour clinic. *Lancet.* 1997;350(9072):187.
- Donaldson AC, Riggio MP, Rolph HJ, Bagg J, Hodge PJ. Clinical examination of subjects with halitosis. *Oral Dis.* 2007;13:63–70.
- Frascella J, Gilbert RD, Fernandez P, Hendler J. Efficacy of a chlorine dioxide-containing mouthrinse in oral malodor. *Compend Contin Educ Dent.* 2000;21:241–54.
- Gerlach RW, Hyde JD, Poore CL, Stevens DP, Witt JJ. Breath effects of three marketed dentifrices: a comparative study evaluating single and cumulative use. *J Clin Dent.* 1998;9:83–8.
- Haraszthy VI, Zambon JJ, Sreenivasan PK, Zambon MM, Gerber D, Rego R, Parker C. Identification of oral bacterial species associated with halitosis. *J Am Dent Assoc.* 2007;138:1113–20.
- Hughes FJ, McNab R. Oral malodour – a review. *Arch Oral Biol.* 2008;53:S1–7.
- Ii H, Imai T, Yaegaki K, Irie K, Ekuni D, Morita M. Oral malodorous compound induces osteoclast differentiation without receptor activator of nuclear factor κ B ligand. *J Periodontol.* 2010;81:1691–7.
- Imai T, Ii H, Yaegaki K, Murata T, Sato T, Kamoda T. Oral malodorous compound inhibits osteoblast proliferation. *J Periodontol.* 2009;80:2028–34.
- John M, Vandana KL. Detection and measurement of oral malodour in periodontitis patients. *Indian J Dent Res.* 2006;17(1):2–6.
- Kazor CE, Mitchell PM, Lee AM, et al. Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients. *J Clin Microbiol.* 2003;41:558–63.
- Lee PP, Mak WY, Newsome P. The aetiology and treatment of oral halitosis: an update. *Hong Kong Med J.* 2004;10:414–8.
- Lee SS, Zhang W, Li Y. Halitosis update: a review of causes, diagnoses, and treatments. *J Calif Dent Assoc.* 2007;35:258–60, 262, 264–8.
- Morita M, Wang HL. Association between oral malodor and adult periodontitis: a review. *J Clin Periodontol.* 2001a;28:813–9.
- Morita M, Wang HL. Relationship between sulcular sulfide level and oral malodor in subjects with periodontal disease. *J Periodontol.* 2001b;72:79–84.
- Morita M, Wang HL. Relationship of sulcular sulfide level to severity of periodontal disease and BANA test. *J Periodontol.* 2001c;72:74–8.
- Murata T, Yaegaki K, Qian W, Herai M, Calenic B, Imai T, Sato T, Tanaka T, Kamoda T, Ii H. Hydrogen sulfide induces apoptosis in epithelial cells derived from human gingiva. *J Breath Res.* 2008;2:017007.
- Outhouse TL, Al-Alawi R, Fedorowicz Z, Keenan JV. Tongue scraping for treating halitosis. *Cochrane Database Syst Rev.* 2006;(2):CD005519. <http://www.ncbi.nlm.nih.gov/pubmed?term=Outhouse%20TL%2C%20Al-Alawi%20R%2C%20Fedorowicz%20Z%2C%20Keenan%20JV.%20Tongue%20scraping%20for%20treating%20halitosis.%20Cochrane%20Database%20Syst%20Rev>
- Pham TAV, Ueno M, Zaitzu T, Takehara S, Shinada K, Lam PH, Kawaguchi Y. Clinical trial of oral malodor treatment in patients with periodontal diseases. *J Periodontol Res.* 2011;46:722–9.
- Quirynen M. Management of oral malodour. *J Clin Periodontol.* 2003;30 Suppl 5:17–8.
- Quirynen M, Zhao H, van Steenberghe D. Review of the treatment strategies for oral malodour. *Clin Oral Invest.* 2002;6:1–10.
- Quirynen M, Zhao H, Soers C, et al. The impact of periodontal therapy and the adjunctive effect of antiseptics on breath odor-related outcome variables: a double-blind randomized study. *J Periodontol.* 2005;76:705–12.

- Ramachandra SS, Mehta DS, Sandesh N, Baliga V, Amarnath J. Periodontal probing systems: a review of available equipment. *Compend Contin Educ Dent*. 2011;32:71–7.
- Ratcliff PA, Johnson PW. The relationship between oral malodor, gingivitis, and periodontitis. A review. *J Periodontol*. 1999;70:485–9.
- Richter JL. Diagnosis and treatment of halitosis. *Compend Contin Educ Dent*. 1996;17:370–86.
- Roldán S, Winkel EG, Herrera D, Sanz M, Van Winkelhoff AJ. The effects of a new mouthrinse containing chlorhexidine, cetylpyridinium chloride and zinc lactate on the microflora of oral halitosis patients: a dual-centre, double-blind placebo-controlled study. *J Clin Periodontol*. 2003;30:427–34.
- Roldán S, Herrera D, O'Connor A, González I, Sanz M. A combined therapeutic approach to manage halitosis: a 3-month prospective case series. *J Periodontol*. 2005;76:1025–33.
- Rosenberg M, Kulkarni GV, Bosy A, McCulloch CA. Reproducibility and sensitivity of oral malodor measurements with a portable sulphide monitor. *J Dent Res*. 1991;70:1436–40.
- Rosenberg M, Gelernter I, Barki M, Bar-Ness R. Day-long reduction of oral malodor by a two-phase oil: water mouthrinse as compared to chlorhexidine and placebo rinses. *J Periodontol*. 1992;63:39–43.
- Schmidt NF, Tarbet WJ. The effect of oral rinses on organoleptic mouth odor ratings and levels of volatile sulfur compounds. *Oral Surg Oral Med Oral Pathol*. 1978;45:876–83.
- Schmidt NF, Missan SR, Cooper AD, Tarbet WJ. The correlation between organoleptic mouth-odor ratings and levels of volatile sulfur compounds. *Oral Surg Oral Med Oral Pathol*. 1978;45:560–7.
- Scully C, Felix DH. Oral medicine – update for the dental practitioner: oral malodour. *Br Dent J*. 2005;199:498–500.
- Scully C, Greenman J. Halitosis (breath odor). *Periodontol* 2000. 2008;48:66–75.
- Seemann R, Kison A, Bizhang M, Zimmer S. Effectiveness of mechanical tongue cleaning on oral levels of volatile sulfur compounds. *J Am Dent Assoc*. 2001;132:1263–7.
- Solis-Gaffar MC, Rustogi KN, Gaffar A. Hydrogen sulfide production from gingival crevicular fluid. *J Periodontol*. 1980;51:603–6.
- Stamou E, Kozlovsky A, Rosenberg M. Association between oral malodour and periodontal disease-related parameters in a population of 71 Israelis. *Oral Dis*. 2005;11 Suppl 1:72–4.
- Takeuchi H, Machigashira M, Yamashita D, Kozono S, Nakajima Y, Miyamoto M, Takeuchi N, Setoguchi T, Noguchi K. The association of periodontal disease with oral malodour in a Japanese population. *Oral Dis*. 2010;16:702–6.
- Tanaka M, Anguri H, Nishida N, Ojima M, Nagata H, Shizukuishi S. Reliability of clinical parameters for predicting the outcome of oral malodor treatment. *J Dent Res*. 2003;82:518–22.
- Tanaka M, Yamamoto Y, Kuboniwa M, et al. Contribution of periodontal pathogens on tongue dorsa analyzed with real-time PCR to oral malodor. *Microbes Infect*. 2004;6:1078–83.
- Tonzetich J. Oral malodour: an indicator of health status and oral cleanliness. *Int Dent J*. 1978;28:309–19.
- Tsai CC, Chou HH, Wu TL, Yang YH, Ho KY, Wu YM, Ho YP. The levels of volatile sulfur compounds in mouth air from patients with chronic periodontitis. *J Periodontol Res*. 2008;43:186–93.
- van den Broek AM, Feenstra L, de Baat C. A review of the current literature on aetiology and measurement methods of halitosis. *J Dent*. 2007;35:627–35.
- van den Broek AM, Feenstra L, de Baat C. A review of the current literature on management of halitosis. *Oral Dis*. 2008;14:30–9.
- Van der Sleen MI, Slot DE, Van Trijffel E, Winkel EG, Van der Weijden GA. Effectiveness of mechanical tongue cleaning on breath odour and tongue coating: a systematic review. *Int J Dent Hyg*. 2010;8:258–68.
- Winkel EG, Roldán S, Van Winkelhoff AJ, Herrera D, Sanz M. Clinical effects of a new mouthrinse containing chlorhexidine, cetylpyridinium chloride and zinc-lactate on oral halitosis. A dual-center, double-blind placebo-controlled study. *J Clin Periodontol*. 2003;30:300–6.
- Wozniak WT. The ADA guidelines on oral malodor products. *Oral Dis*. 2005;11(Suppl) 1:7–9.
- Yaegaki K, Coil JM. Examination, classification, and treatment of halitosis; clinical perspectives. *J Can Dent Assoc*. 2000a;66:257–61.
- Yaegaki K, Coil JM. Genuine halitosis, pseudo-halitosis, and halitophobia: classification, diagnosis, and treatment. *Compend Contin Educ Dent*. 2000b;21(10A): 880–6, 888–9.
- Yaegaki K, Sanada K. Biochemical and clinical factors influencing oral malodor in periodontal patients. *J Periodontol*. 1992a;63:783–9.
- Yaegaki K, Sanada K. Volatile sulfur compounds in mouth air from clinically healthy subjects and patients with periodontal disease. *J Periodontol Res*. 1992b;27:233–8.

17.1 Periodontitis Is a Complex Disease

What if you had a dresser drawer filled with thousands of puzzle pieces mixed together from an unknown number of jigsaw puzzles, and you wanted to solve each puzzle separately? How would you begin to sort this out? This puzzle drawer is analogous to what many researchers face when trying to discover the underlying genetic and environmental causes linked to complex forms of human disease (Craig 2008).

17.1.1 What Is a Complex Disease?

For the most part, complex diseases are caused by a combination of genetic, environmental, and lifestyle factors, most of which have not yet been identified. The vast majority of diseases fall into this category, including several congenital defects and a number of adult-onset diseases. Some examples include Alzheimer's disease, scleroderma, asthma, Parkinson's disease, multiple sclerosis, osteoporosis, connective tissue diseases, kidney diseases, autoimmune diseases, and many more (Craig 2008).

Periodontitis is a complex disease that develops in a limited subset of humans. About 10% of the population will develop severe forms of destructive periodontal disease (**Fig. 17.1**) (Schäfer et al. 2011). Smoking increases the severity of both

chronic periodontitis and aggressive periodontitis (because of their young age, the effect on the localized form of aggressive periodontitis is not clear). Oral hygiene, as assessed by plaque levels, is directly associated with disease severity in both entities, although in the localized form of aggressive periodontitis, studies have suggested no correlation between plaque levels and the presence of disease. Plaque level is a notoriously fickle indicator of oral hygiene because of the day-to-day variations resulting from individual behavioural patterns. Psychological factors seem to play a role in both chronic periodontitis and aggressive periodontitis, although our knowledge regarding this issue is severely lacking (Stabholz et al. 2010).

Systemic diseases, as defined by the 1999 classification, cannot be considered as risk factors for either chronic periodontitis or aggressive periodontitis. Cases of severe periodontitis that can be shown to be caused by profound host-modifying effects (e.g., major immunosuppression) of a known systemic disease are properly categorized as 'periodontitis as a manifestation of systemic diseases'. However, systemic diseases that can cause subtle perturbations in host susceptibility to infections (e.g., diabetes mellitus) can alter the clinical course of chronic periodontitis and aggressive periodontitis (Stabholz et al. 2010).

Scientists now know that **complex diseases do not obey the standard Mendelian patterns of inheritance**. Although we inherit genes associated with these diseases, genetic factors represent

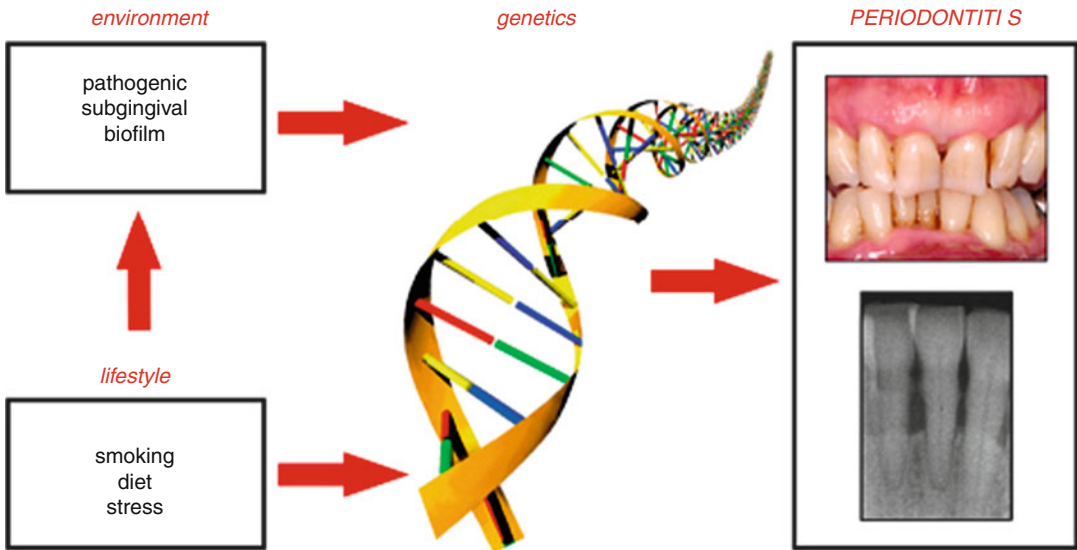


Fig. 17.1 Periodontitis is a complex disease. The modifying disease genes play a central role and determine susceptibility to periodontitis. With this intrinsic characteristic, lifestyle factors pave the way to the disease, which is eventually initiated by an infection (environmental factors). Both have an individual as well as an additive effect on the (patho)physiological response of the patient,

which is, in the final step, determined by the individual genetic predisposition. Therefore, in the complexity of periodontitis, gene–gene interactions as well as gene–environmental, gene–lifestyle, and environmental–lifestyle interactions play a role in the development of the phenotype (Schäfer et al. 2011. Reprinted with permission from John Wiley & Sons. Inc.)

only part of the risk associated with complex disease phenotypes. A genetic predisposition means that an individual has a genetic susceptibility to developing a certain disease, but this does not mean that a person harbouring a genetic tendency is destined to develop the disease. The actual development of the disease phenotype depends in large part on a person's environment and lifestyle. While we cannot change our genes, we can alter our lifestyle and environment to prevent or delay the onset of such a disorder. Indeed, the interplay between genetic and environmental factors in complex disease continues to challenge researchers (Craig 2008; Dempfle et al. 2008).

Considering gene–environment interactions can improve our understanding of the causes of complex disease, and it can also assist researchers in developing targeted therapies. Scientists now understand that gene products and the by-products of environmental insult often interact at the molecular level. Rather than studying genetic and environmental factors separately, researchers

are now studying how genes and environmental factors interact with one another. By looking at the whole picture, researchers can identify genetic risk factors, which may in turn be modified in an environment-specific manner (Craig 2008; Dempfle et al. 2008).

17.1.2 The Human Genome Project and New Approaches in Gene Searching

Genomics is the field of study that seeks to understand the role of the genome in the development, function, and organization of cell types into higher-order structures. The completion of the Human Genome Project has changed how researchers approach complex diseases by revealing new insights as to the genetic pathogenesis of disease. For example, scientists have discovered that the human genome sequence possesses numerous variations, of which single

nucleotide polymorphisms (SNPs) are the most prevalent. As their name implies, SNPs are single DNA base-pair changes, located throughout our genome, that frequently vary from one person to another (Craig 2008; Kinane and Hart 2003).

Genome-wide association (GWA) studies are studies in which a dense array of genetic markers, which captures a substantial proportion of common variation in genome sequence, is typed in a set of DNA samples that are informative for a trait of interest. The aim is to map susceptibility effects through the detection of associations between genotype frequency and trait status (McCarthy et al. 2008).

The first wave of large-scale, high-density genome-wide association (GWA) studies has improved our understanding of the genetic basis of many complex traits. For several diseases, including type 1 and type 2 diabetes, inflammatory bowel disease, prostate cancer, and breast cancer, there has been rapid expansion in the numbers of loci implicated in predisposition. For others, such as asthma, coronary heart disease, and atrial fibrillation, fewer novel loci have been found, although opportunities for mechanistic insights are equally promising. Several common variants influencing important continuous traits, such as lipids, height, and fat mass, have also been found. An updated list of published GWA studies can be found at the National Cancer Institute (NCI)–National Human Genome Research Institute (NHGRI)’s catalogue of published genome-wide association studies (McCarthy et al. 2008 and references therein).

17.1.3 Describing the Gene–Environment Interactions

Even with well-designed studies, there are many ways of declaring ‘success’ in the search for interactions. This is largely because of variability in the qualitative and statistical models of interaction and the difficulty of assessing biological plausibility (either *a priori*—i.e., when trying to prioritize prior probabilities in these analyses—or once an interaction has been observed). These factors are described below (Hunter 2005).

17.1.3.1 Qualitative Models

In the simplest case of dichotomous genotype (such as carriers vs. non-carriers of a gene variant) and dichotomous exposure (e.g., exposed vs. non-exposed), the four possible combinations of genotype and exposure can be displayed in a 2×4 table, and the relative risks can be shown in a graph such as that shown in Fig. 17.2a. However, even in this simplest case, there are several models for describing interactions between genetic susceptibility and environmental exposures in different diseases. The possibilities are more numerous if there are many categories of environmental exposure (e.g., three or more categories of exposure; see Fig. 17.2b) and/or many genetic categories (as in the case of three genotypes for a biallelic system) or different genetic models (recessive, co-dominant, and dominant) (Table 17.1). So, many departures from the null result—where the risk of disease is the same in all cross-classified categories of exposure and genotype—might be compatible with the overall hypothesis of gene–environment interaction (Hunter 2005).

The ‘multiple comparison’ problem that is inherent in examining thousands or even hundreds of thousands of SNPs in association studies is relatively familiar. For gene–environment interactions, however, we face a comparison problem that arises from a model involving multiple genes, multiple exposures, and multiple interactions. In addition to using statistical approaches to control the false-positive rate, the reproducibility of gene–environment interactions across two or more studies will be crucial (Hunter 2005).

17.1.3.2 Statistical Models

In addition to the many qualitative models of interaction described, there are several methods of assessing the statistical significance of interactions. In the simplest case of dichotomous environmental exposure and genotype, perhaps the most commonly used procedure is to test departure from the multiplicative model of interaction. This involves testing whether the relative risk for joint exposure (cell d in Fig. 17.2a) is statistically significantly

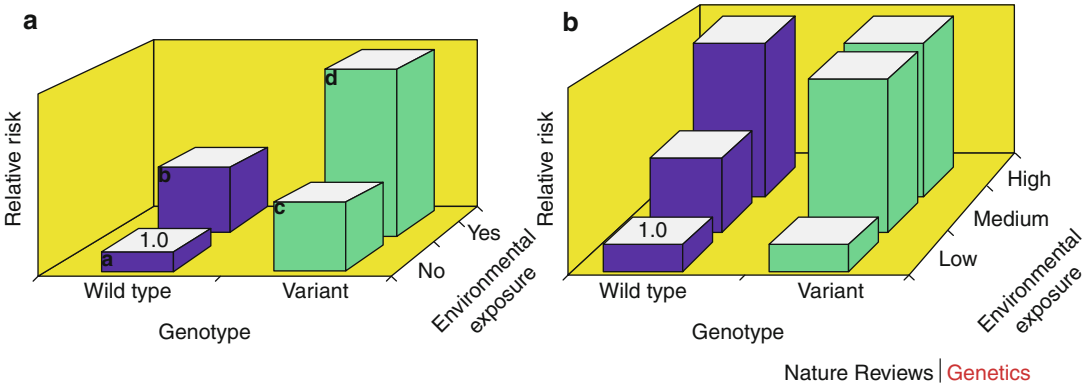


Fig. 17.2 Models of gene–environment interactions. (a) In the most simplified example of a dichotomous genotype (e.g., carriers vs. non-carriers of an allele corresponding to a dominant trait), and dichotomous exposure (e.g., ‘exposed’ vs. ‘non-exposed’), three categories of joint exposure can be compared with a reference category (for which the relative risk is, by definition, 1.0). Using this simple scheme, **Table 17.1** shows the different patterns of risk that are observed in some diseases in which inherited susceptibility clearly interacts with environmental exposures to jointly determine disease risk. In the example shown here, the relative risk of developing a disease is much greater in individuals who are both genetically susceptible to the condition and have been exposed to the

environmental variable (cell d), than in individuals who carry the wild-type genotype and are not exposed to the environmental variable (cell a), or who are either only exposed to the environment or genetically susceptible (cells b and c, respectively). (b) In the slightly more complex situation in which there are three categories of exposure, it has been proposed that genetically susceptible individuals could be at risk of disease at lower levels of exposure; in this model, the difference in risk between genotypes among individuals at the medium level of exposure is the only indication of an interaction (Hunter 2005. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

Table 17.1 Some patterns of relative risk in gene–environment interactions (Hunter 2005) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

The table shows just three examples of different patterns of relative risk for three classical genetic diseases that have an environmental component, assuming dichotomous genetic susceptibility and environmental exposure. In the first example, xeroderma pigmentosum (XP), exposure to ultraviolet light increases the risk of developing skin cancer in non-carriers of XP mutations, but the combination of these mutations and exposure to ultraviolet light vastly increases the risk of skin cancer. In theory, if individuals with XP mutations completely avoid ultraviolet light their risk of skin cancer becomes close to the background risk.

The example in the second column is that of phenylketonuria (PKU); only individuals with recessive mutations in the causative gene (phenylalanine hydroxylase) that are exposed to phenylalanine in the diet are susceptible to PKU.

In the third column, exemplified by a deficiency in the α -1 antitrysin gene, both non-smokers that are at genetic risk and smokers that are not at genetic risk have an increased risk of developing emphysema, and the combination (smokers that are at genetic risk) is associated with the highest risk.

There are many other patterns of gene–environment interactions, including ‘protective’ alleles and exposures.

Gene variant	Environmental exposure	Relative risk (XP)	Relative risk (PKU)	Relative risk (emphysema)
Absent	Absent	1.0	1.0	1.0
Present	Absent	~1.0	1.0	Modest
Absent	Present	Modest	1.0	Modest
Present	Present	Very high	Very high	High

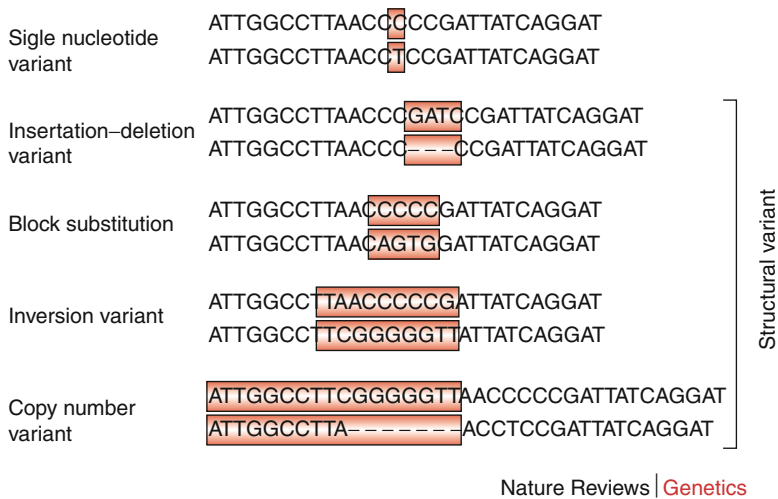


Fig. 17.3 *Classes of human genetic variants.* The nomenclature used to describe the various types of structural variants is not yet standard 121. Here, the terminology used aims to describe the nucleotide composition of the variant and distinguish it from other types of variants. Single nucleotide variants are DNA sequence variations in which a single nucleotide (A, T, G, or C) is altered. Insertion-deletion variants (indels) occur when one or more base pairs are present in some genomes but absent in others. They are generally composed of only a few bases but can be greater than 80 kb in length 11. Block substitutions describe cases in which a string of adjacent nucleotides varies between two genomes. An inversion

variant is one in which the order of the base pairs is reversed in a defined section of a chromosome. A well-characterized inversion variant that has been described in humans involves a section of chromosome 17 in which an ~900-kb interval is in the reverse order in approximately 20% of individuals with Northern European ancestry 122. Copy number variants occur when identical or nearly identical sequences are repeated in some chromosomes but not others. The largest copy number variant identified in the Venter genome 11 was almost 2 MB in length (Frazer et al. 2009. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

greater (‘supermultiplicative’) or smaller (‘submultiplicative’) than would be expected by multiplying the relative risks for environmental exposure or genetic predisposition alone (i.e., multiplying the relative risks of cells b and c in **Fig. 17.2a**). Another commonly used test for interaction uses rate differences rather than relative risks and proposes that the joint effect of genes and the environment is different from the expectation that the incidence rate in cell d in **Fig. 17.2a** is described by adding the rates in cells b and c in **Fig. 17.2a**. This additive model is often said to be of greater relevance to assessing the public health impact of an interaction. Again, the option to use either of these two models adds further potential for multiple comparisons to the statistical analysis of gene-environment interactions (Hunter 2005).

17.1.4 Classes of Human Genetic Variation

Human genetic variants are typically referred to as either common or rare, to denote the frequency of the minor allele in the human population. Common variants are synonymous with polymorphisms, defined as genetic variants with a minor allele frequency (MAF) of at least one per cent in the population, whereas rare variants have a MAF of less than 1%. Genetic variants are also discussed in terms of their nucleotide composition. In the broadest sense, variants in the human genome can be divided into two different nucleotide composition classes: single nucleotide variants and structural variants10 (**Fig. 17.3**) (Frazer et al. 2009).

17.1.5 Study Designs for Genetic–Environment ($G \times E$) Interactions

Common family- and population-based designs for association studies can be extended for $G \times E$ interaction. **Table 17.2** lists different designs with their respective advantages and disadvantages and research situations in which such a design would be suitable. **Family-based designs** protect against bias due to population stratification with both differential exposure and genotype distribution in subgroups. In **population-based designs**, data on a quantitative trait or a disease phenotype are collected from unrelated individuals, either prospectively (cohort) or retrospectively (case–control). If a large prospective cohort exists, a nested case–control study can reduce selection and possibly stratification biases and be a good compromise regarding cost and efficiency. It was also argued that **case–control studies** are more feasible and cost-efficient than cohort studies for modest disease risks and that exposure misclassification bias is not a serious threat in the case of $G \times E$ interactions. If the interest is limited to $G \times E$ interaction, the special ‘**case-only**’ **design** exists that has the practical advantage that no controls need to be collected. This design is based on the assumption that genotype and environmental exposure are independent in the population that the case sample is drawn from, so that exposure should not differ among subgroups defined by genotype. Since, in the presence of a $G \times E$ interaction, specific combinations of genotypes and exposure lead to increased risk of disease and thus are more prevalent among cases, differences in exposure will be observable between genotype groups in cases. Because of the independence assumption, the case-only design is more efficient than the traditional case–control design, but this assumption is not assessable in the case sample alone. Therefore, the design is prone to bias and confounding, especially if there is exposure misclassification (keeping in mind that especially lifetime environmental exposures are not as accurately measurable as genotypes). Another drawback is that although estimation of the $G \times E$ interaction is possible, the estima-

tion of the joint effect of exposure and genotype is impossible even though the latter usually is of greater importance for the public health aspect of a $G \times E$ investigation. As a consequence, the practical applicability of this design is limited and it is rarely applied (Dempfle et al. 2008).

Two special, non-standard applications of $G \times E$ interactions occur in infectious disease and pharmacogenetic studies. In infectious diseases, only individuals exposed to the infectious agent can contract the disease; thus, the environmental factor is a necessary causal factor. Similarly, some pharmacogenetic studies for licenced drugs aim at identifying individuals at risk for serious side effects or increased efficacy by exclusively including drug-treated patients. In this design, it is impossible to distinguish between genetic effects and $G \times E$ interaction. More suitable is a design that includes pharmacogenetic aspects in randomized clinical trials by giving placebo or active drug stratified according to genotype (Dempfle et al. 2008).

17.2 Polygenetic Model for Periodontitis

In general, the genetic basis of diseases is divided into two major groups. One is the result of a gene defect or a mutation, which induces common disease phenotypes, but is not influenced by the host’s genetic background. The other is that a disease develops as the result of a cumulative effect of multiple genes, each of which by itself cannot induce the particular phenotype. The former indicates an almost qualitative phenotype, while the latter indicates a quantitative phenotype, which is not dependent on Mendelian inheritance. This system was designated ‘polygenic inheritance’ by K. Mather in 1949. Each polygene may distribute into the population of a species through an evolutionary process, influenced by genetic drift, and taking advantage of escape from natural selection (Nose 2011).

For example, collagen disease is under the control of polygenes according to the law proposed by Mather and involves polymorphic genes and

Table 17.2 Study designs for genetic association studies that can include G×E interactions with their main advantages and disadvantages and the situations in which these designs are most suitable (Dempfle et al. 2008) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

Study design	Main advantage	Disadvantage	In which situation is this design most suitable?
<i>Family-based designs</i>			
Trio design: case and both parents, evaluate transmission disequilibrium, depending on exposure	Protects against confounding due to population stratification	High ascertainment costs; often impossible for late onset diseases	Early onset diseases when population stratification is a concern
Sib design: case and unaffected siblings	Protects against population stratification	Potentially difficult to recruit enough suitable families	Late onset diseases when population stratification is a concern
<i>Population-based designs</i>			
Cohort: compare disease frequency between genotype-exposure groups	Reduces selection and stratification bias compared to case-control design; prospective measurement of exposure	Expensive and time consuming; investigation of very rare diseases may be impossible with realistic sample sizes	Very reliable results due to low risk of biases, therefore suitable for confirmatory studies; useful for common diseases; can use existing cohorts with additional DNA collection
Case-control (retrospective): compare exposure rate and genotype frequencies between cases and controls	Simple, relatively inexpensive and less time consuming compared to cohort studies	Possible stratification bias due to population stratification. Higher risk of measurement error in exposure	Only reasonable population-based design for very rare diseases; very suitable for first exploratory studies
Case-control (nested, prospective)	Less expensive than full cohort design, reduced biases compared to retrospective case-control design		For rare or common diseases in existing cohorts; confirmatory studies
Case-only: evaluate differences in exposure between genotype groups in cases	Simple, relatively inexpensive and less time consuming compared to cohort studies; no controls needed; high power for G×E interaction test	Assumption of independence of G and E not assessable in the case-only sample, therefore prone to bias and confounding; joint effect of G and E cannot be estimated	Fast, uncomplicated exploratory studies, results need to be confirmed with other study designs
Clinical trial, randomization stratified by genotype	Gives reliable results of clinical relevance of a presumed gene-by-drug interaction	Increased sample size and higher cost over traditional clinical trial	If preliminary studies suggest differential drug effects by genotype; basis for personalized medicine
Exposed-only for infectious diseases	No need to include a presumably unexposed group		Infectious disease genetics

also a mutant gene encoding lymphoproliferation in the case of the MRL/lpr mice. Polygenes may quantitatively regulate the cascade reactions extending to the development of vasculitis and other collagen disease with a regular variation of the pathological phenotypes based on their combinations (Nose 2011). Recently, polygenetic models were developed to explain the susceptibility profile for glaucoma (Jia et al. 2009), asthma (Lee et al. 2008), diabetes (Weedon et al. 2006), epilepsies (Dibbens et al. 2007), stroke (Shen et al. 2007), and cancer (Listgarten et al. 2004).

It has been estimated that more than 20 genes and hundreds of polymorphisms of these genes are associated with chronic periodontitis and aggressive periodontitis (Dumitrescu and Konbayashi 2009). Of these, approximately 10 genes have been comprehensively studied, including the interleukin-1 (IL-1), interleukin-4, interleukin-6, tumor necrosis factor α (TNF α), vitamin D receptor (VDR), the Fc γ receptor (Fc γ R), matrix metalloproteinase, Toll-like receptors, cyclo-oxygenase-2 (COX-2), and C-reactive protein. However, the frequency distribution of the candidate gene polymorphisms does not correlate with the prevalence of periodontitis. It is likely that a certain gene may have a minor effect on disease onset and pathogenesis, and then other genes may interact with each other, with no single gene having a dominant effect. The odds ratios in initial studies on polymorphisms for disease associations are usually less than 3.0. However, these figures are rarely replicated in later studies, and thus it is difficult to estimate population-wide effects of genetic risk factors in chronic disease (Zhang et al. 2011).

Generally, efforts to find major risk factors for complex, polygenic diseases using GWA studies only have been moderately successful. A widespread explanation is that complex diseases, unlike rare and monogenic diseases, may be caused by multiple individual susceptibility alleles with low effect sizes. The modest contributions of rare risk alleles combined with massive statistical chance correlations in the genome-wide association (GWA) analyses require large cohorts and refined single nucleotide polymorphism (SNP) selection. A common strategy to alleviate the problems of detecting common variants is to expand the study cohorts. Crohn's disease and lipid levels are examples where large collaborative meta-analysis has uncovered

between 20% and 50% of the genetic variance (Pers et al. 2011 and references therein). Pers et al. (2011) proposed a widely applicable approach to identify susceptibility genes based on meta-analysis of heterogeneous molecular data sets, which is complementary to GWA-based meta-analyses that combine data of the same type. This presented flexible method augments modestly or underpowered GWA data and prioritizes the genome in relation to the phenotype of interest by integrating potentially complementary evidence layers of heterogeneous data sources for a given risk phenotype or disease.

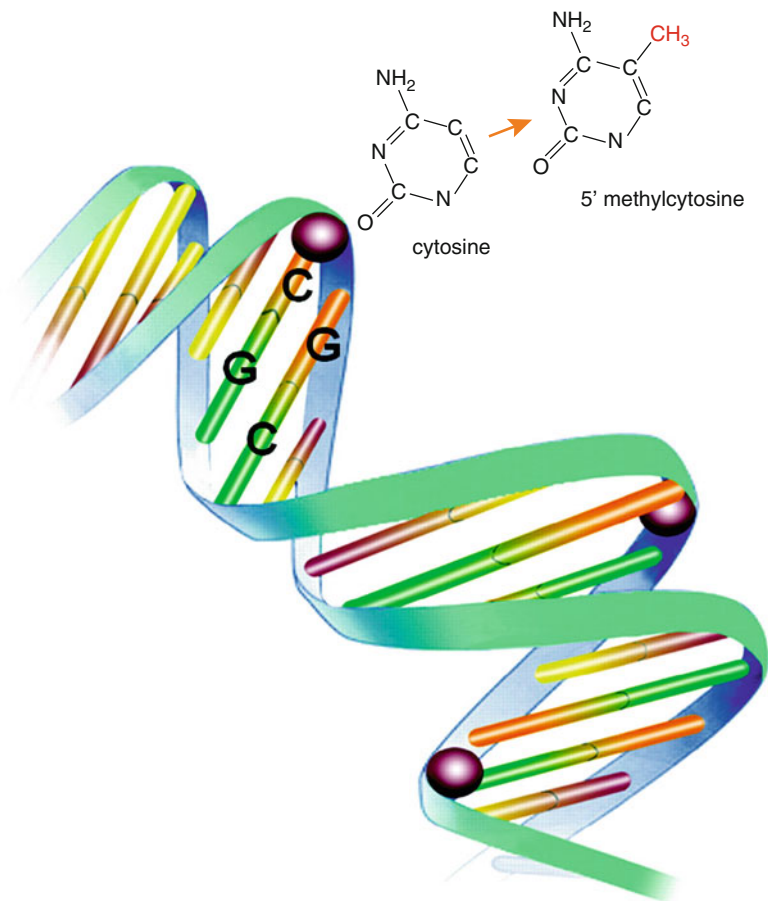
They include:

1. SNP to phenotype associations from GWA studies which represent a rapidly growing resource of unbiased genome-wide associations of common risk alleles.
2. Interacting pairs of candidate proteins and proteins encoded by known phenotype susceptibility genes—a type of data which also targets rare alleles. Two proteins involved in the same biological (dys)function often interact.
3. Data from linkage studies capturing co-segregation of chromosomal regions and disease-specific phenotypes in families is another methodology complementary to GWA analyses capable of identifying regions harbouring rare disease-specific variants.
4. Quantitative data on disease similarities, which may add information that cross normal disease definition barriers.
5. Gene expression levels may be affected directly or indirectly by polymorphisms associated with disease. Differential expression between cases and controls may also add important tissue-specific information.

The strategy allows for the integration of complementary data sources in a single meta-analysis, leading to a prioritization of protein-coding genes from the entire genome in relation to one particular indication. The data sources can be perceived as layers that are collapsed into an integrative meta-layer, providing an informed selection of new candidates. The integration provides a list of high-ranking candidate genes with robust support from the different evidence layers, where a small number of genes subsequently can be subjected to further experimental analysis (Pers et al. 2011).

While meta-analyses may provide more meaningful sound evidence, very few such studies have been

Fig. 17.4 Methylation of DNA occurs at cytosine residues when present as CG dinucleotides. Methylation occurs by the addition of a methyl group at the 5' site of cytosine (depicted as *shaded sphere*) (Barros and Offenbacher 2009. Reprinted with permission from SAGE Publications)



performed (Zhang et al. 2011; Nikolopoulos et al. 2008; Thakkinian et al. 2005). The first quantitative evaluation of polymorphisms and periodontitis showed that the IL-1 β +3954T variant is an indicative marker for Asian populations and has a synergistic effect with IL-1 α -889T. These two variants may be the best candidate genes for a polygenetic model of periodontitis in Asian populations. However, this study covered only six polymorphisms in four genes. More polymorphisms need to be subjected to meta-analyses in order to develop a more accurate polygenetic model (Zhang et al. 2011).

17.3 Epigenetics and Periodontal Disease

Epigenetics is a relatively new concept in periodontitis research that may uncover the missing link between genetics, disease, and environment (Hirst and Marra 2009; Laine et al. 2012).

Epigenetic regulation is the mechanism by which gene function is selectively activated or inactivated in the cells. It provides higher-ordered and more specified genetic information, compared with the whole genome itself (Nakao 2001). There are two major forms of epigenetic regulation: posttranslational modification of DNA-associated histone proteins in chromatin and methylation of DNA. These forms are regulated by distinct but coupled pathways. Notably, histone Lys acetylation by histone acetyltransferase and deacetylation by histone deacetylases play a crucial role in on-off regulation of gene expression (Fig. 17.4) (Imai and Ochiai 2011).

In 2008, the NIH Roadmap Epigenomics Mapping Consortium (<http://www.roadmap-epigenomics.org/>) was launched with the goal of producing a public resource of human epigenomic data to catalyse basic biology and disease-oriented research (Fig. 17.5). The consortium leverages experimental pipelines built

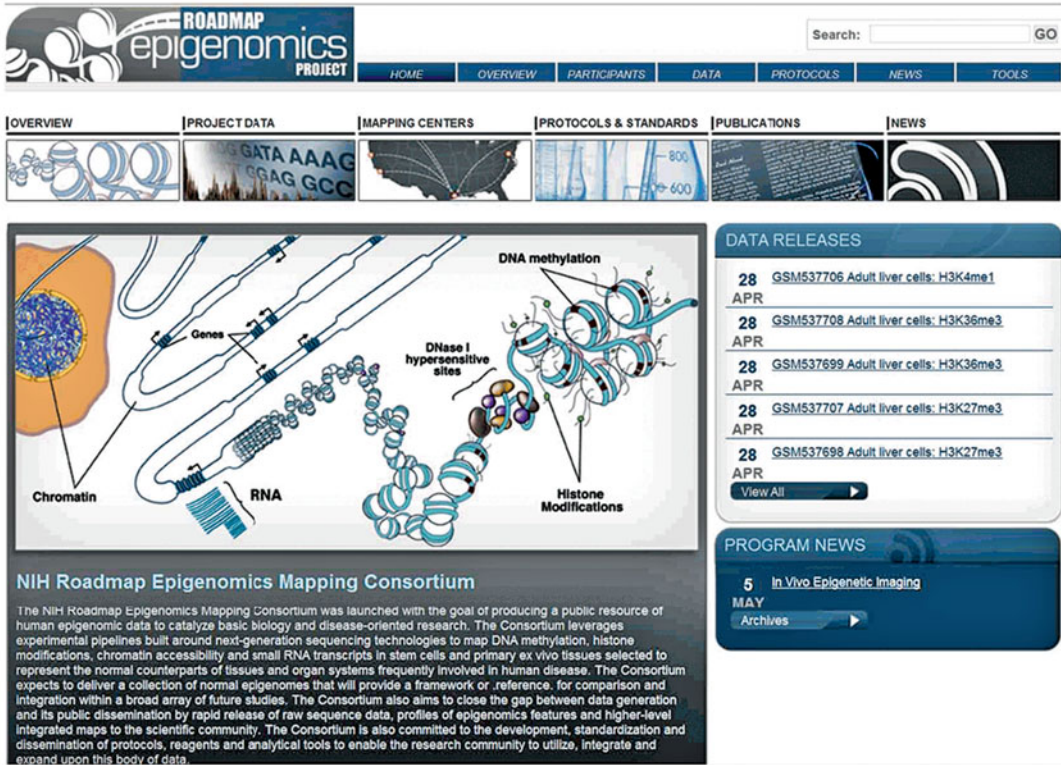


Fig. 17.5 Portal for the NIH Roadmap Epigenomics Mapping Consortium (Bernstein et al. 2010. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

around next-generation sequencing technologies to map DNA methylation, histone modifications, chromatin accessibility, and RNA transcripts in stem cells and primary *ex vivo* tissues selected to represent the normal counterparts of tissues and organ systems frequently involved in human disease. The mapping of such normal epigenomes is being undertaken by four Epigenomics Mapping Centers and supported by a Data Analysis and Coordinating Center, which collectively coordinate experimental and analytical efforts to maximize consistency, data quality, and overall coverage of the epigenomic landscape (Bernstein et al. 2010).

Studies have demonstrated that epigenetic events are able to influence the production of cytokines, contributing to the development of inflammatory diseases. The epigenetic process, by inducing a change in cytokine profile, may subsequently influence the pathogenesis and

determine the outcome of many infectious diseases. These findings may have relevance for inflammatory diseases in which the expression of cytokines is unregulated. A complex network of pro- and anti-inflammatory cytokines acts on the inflamed periodontal tissues, such as IL-1 beta, IL-1 alpha, TNF-alpha, IL-6, IL-10, and IL-4, and an overexpression of inflammatory cytokines has been demonstrated in individuals with periodontal disease (Gomez et al. 2009). Preliminary findings suggested that IL-6 gene is hypomethylated in tissues of individuals with periodontal disease compared to control samples, suggesting an overexpression of this cytokine in inflamed tissues (Offenbacher et al. 2008). Additionally, the overexpression of IL-6 might exert an epigenetic influence in cells by regulating the DNMT gene or by maintaining its methylation status, as previously shown in *in vitro* studies (Stenvinkel et al. 2007; Hodge et al. 2005). These findings are

important since IL-6 is a key cytokine involved in bone resorption and has been detected in high levels in individuals with severe periodontitis (Moreira et al. 2007; Gomez et al. 2009).

Recent evidence also indicates that certain orange and red complex bacteria (especially *C. rectus*) can cause epigenetic changes in cells and tissues (Barros and Offenbacher 2009). The organisms gain significant survival advantages as these changes shut down genes involved in local defences and healing (facilitating colonization and dissemination), in the regulation of vascular endothelial function (decreasing tissue perfusion and increasing inflammatory damage), and in host metabolism (providing bacteria with more carbohydrates to fuel growth and reproduction) (Iacopino 2010).

It has been also showed that epigenetic regulation is involved in the establishment and maintenance of HIV-1 latency and in the reactivation of HIV-1 by periodontopathic bacteria. The effect of histone modification on transcriptional regulation and the contribution thereof to the regulation of HIV-1 gene expression during the lytic and latent stages of HIV-1 infection, as well on the mechanisms by which periodontal diseases may accelerate AIDS progression in infected individuals as a new systemic disease caused by periodontitis, were also investigated (Imai and Ochiai 2011).

Considering the findings above, epigenetic factors may interfere in the response of host to bacterial challenge and may contribute to the phenotype of periodontitis (Gomez et al. 2009).

Studies evaluating the occurrence of epigenetic events in periodontal disease must be performed with caution. Periodontal disease has a multifactorial nature, and several factors may influence both the epigenetic status and the periodontal conditions of individuals, such as smoking and age. The inclusion of a homogeneous population from the same geographical area, with similar socio-economic status and matched for confounding factors must be selected. A study with these characteristics should also consider the clinical parameters before associating it with epigenetic events. It can be envisaged that new diagnostic or therapeutic tools could be developed by the manipulation of the host and/or bacterial epigenome (Gomez et al. 2009).

17.4 Challenges for Periodontitis Genetics Studies

The candidate gene approach tries to identify one allele of a gene that is more frequently seen in subjects with the disease than in subjects without the disease. Candidate genes are chosen on the basis of their known or presumed functions that are thought to have some plausible role in the disease. There are three types of candidate genes: functional candidate genes, positional candidate genes, and expressional candidate genes. Functional candidate genes are derived from an existing knowledge of the phenotype and the potential function of the gene involved after clinical or physiological studies of affected individuals. Positional candidate genes are based on the involvement of the gene to a marked location after genetic linkage analyses. Expressional candidate genes are determined through differences in gene expression using microarrays (Yoshie et al. 2007; Hodge 1993).

For several genes, which have been individually sequenced for association with periodontitis, we see a scattered picture from different studies of varied populations and ethnicities. To produce scientifically sound and meaningful disease-association studies, some issues and concerns should be addressed (Yoshie et al. 2007).

17.4.1 Ethnic Heterogeneity

Ethnic and racial differences in disease-susceptibility gene polymorphisms have been frequently reported. This is especially the case when Asian populations are compared with Caucasian populations. Thus, specific gene polymorphisms strongly associated with the disease and showing a high distribution in one ethnic group may not be relevant in another ethnic group because of their low frequency. Therefore, high-frequency population-specific alleles are particularly useful for mapping genes responsible for disease susceptibility. So far, the frequency in various populations of majority gene polymorphisms is not available. It is important that subjects matched by ethnic and geographic origin are chosen in periodontitis association studies in order to control

misleading results (Zhang et al. 2011). For example, polymorphisms in NOD2 gene are causative for Crohn's disease in Caucasian populations but do not play a role in Asians, although the clinical presentation of the disease is identical. The data show that R allele carriage frequencies for the IL1A -889, IL1B +3954, IL1RN VNTR, IL6 -174, IL10 -1087, and VDR *TaqI* and *BsmI* polymorphisms are not or hardly polymorphic in Asian populations in contrast to Caucasian populations, emphasizing the importance of accurate definition of the ethnic background of study cohorts (Laine et al. 2012).

17.4.2 Clinical Classification of Periodontal Disease

Classifying periodontal diseases has been a long-standing dilemma largely influenced by paradigms that reflect the understanding of the nature of periodontal diseases during a given historical period (Yoshie et al. 2005, 2007). Two forms of periodontal disease, that is, chronic periodontitis and aggressive periodontitis, have been proposed (Armitage 1999). However, this classification is clearly unsatisfactory because differences exist within these groups in terms of the extent, severity, and age of onset of disease, and both forms of disease probably include periodontal diseases (Laine et al. 2012). It is the experience of many clinicians that in many instances, the differential diagnosis between these two diseases is often difficult, if not impossible, to make. There is an overlapping 'grey area' in which their definition (according to the 1999 definition) does not allow for a clear-cut diagnosis and therefore raises the question as to how different these two entities really are. For this reason, it is important to review the literature relating to the various risk factors, of which some are considered to be discriminatory for the two disease entities (Stabholz et al. 2010). Inaccuracy in disease classification and definition make interpretation and replication of the case-control studies difficult. Strict disease classifications should be used during selection of the patients (Laine et al. 2012).

One of the main risk factors for aggressive periodontitis is the familial aggregation associated with inherited genetic traits. There is strong evidence

showing familial aggregation in young patients with early expression of aggressive periodontal disease. However, familial aggregation is not always detected, and in many cases, the family members are not available for examination. The efforts that have been carried out in the last 20 years to identify specific gene variations involved in the diseased have not been conclusive. In addition, studies suggest that chronic periodontitis also has a genetic basis, making the discrimination between the two diseases, on a genetic basis, very difficult. The evidence for genetic influences in both diseases exists, but its effect on disease expression is not understood (Stabholz et al. 2010).

17.4.3 Sample Size of the Study Subjects

Research on genetic polymorphisms has only had limited success in identifying significant and reproducible genetic factors for susceptibility to aggressive periodontitis and chronic periodontitis. Taking together the data published on gene polymorphisms in aggressive periodontitis and chronic periodontitis, there are differences among the various studies for the R allele carriage rates, even if the study populations are of the same ethnic background. Nevertheless, there is some evidence that polymorphisms in the IL1B, IL1RN, FcγRIIb, VDR, and TLR4 genes may be associated with aggressive periodontitis, and polymorphisms in the IL1B, IL1RN, IL6, IL10, VDR, CD14, TLR4, and MMP1 genes may be associated with chronic periodontitis susceptibility as a single genetic factor in certain populations. However, most of the results have not been replicated in study cohorts with 100 or more individuals (Laine et al. 2012).

Variants that are associated with complex diseases increase the disease risk rather modestly. The odds ratio (OR) of such a variant usually does not exceed 1.3. Accordingly, to identify a variant with an OR of 1.3 and of an allele frequency of 20% in the general population, 1,000 well-defined cases are necessary to reach the necessary statistical power of 0.8. To match these cases, at least the same number of controls is necessary, ideally twice as many (Schäfer et al. 2011).

In genetic association studies, the genetic effect size for associated markers tends to be overestimated as a consequence of the winner's curse. This bias is due to the strong positive correlation between the association test statistic and the estimator of the genetic effect and the focus of investigators on markers that show statistically significant evidence of association. Analytic calculations that quantify the impact of the winner's curse in large-scale genetic association studies were presented and confirm that in realistic situations, it can result in substantial overestimation of the true genetic effect as measured by the case-control allele frequency difference or the

corresponding odds ratio. A maximum likelihood estimators were proposed that correct for the typical focus on statistically significant results and demonstrate that this estimator results in reduced absolute bias compared to the naive uncorrected estimator when study power is low or moderate ($>60\%$), a range that is typical for most large-scale genetic association studies, and similar absolute bias when power is high. Software that carries out this analysis for case-control data is available at <http://csg.sph.umich.edu/boehnke/winner> (Xiao and Boehnke 2009, 2011).

Some illustrative examples for association studies of a candidate gene are shown in **Fig. 17.6**,

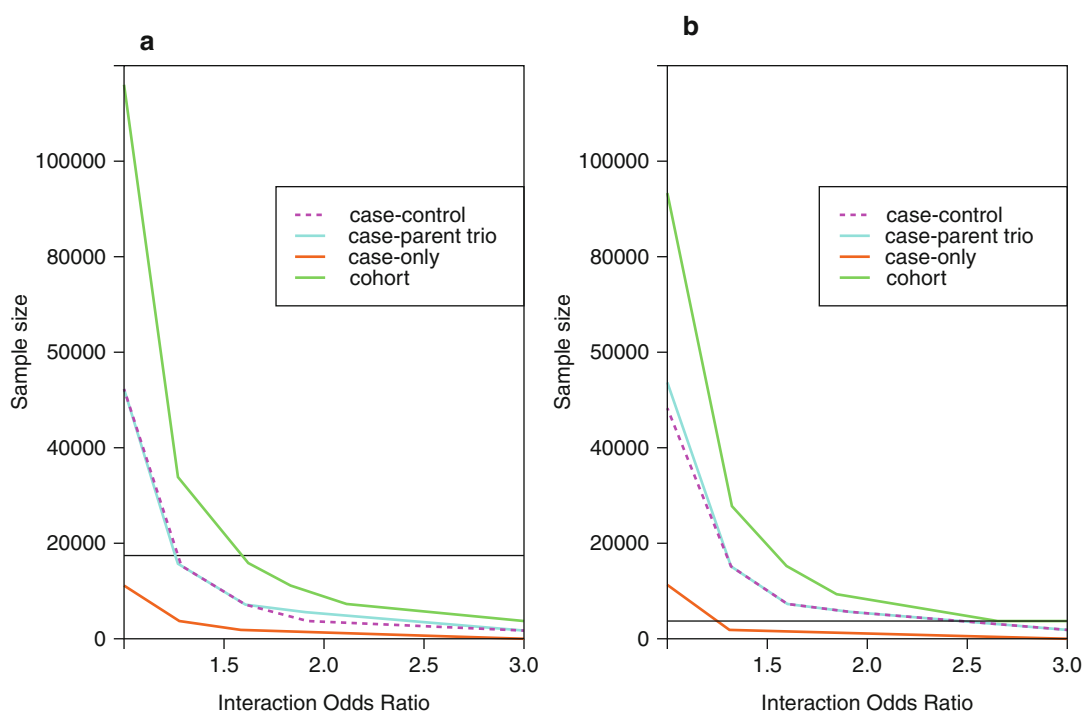


Fig. 17.6 Sample size requirements for 80% power to detect a gene-environment ($G \times E$) interaction for different study designs depending on the strength of the interaction. Sample sizes for case-control, case-parent trio, and case-only designs were calculated using Quanto⁵⁹ (<http://hydra.usc.edu/gxe>), assuming an analysis by (conditional) logistic regression. For the cohort design, sample sizes are estimated using Power⁵⁸ (<http://dcegqa.cancer.gov/bb/tools/power>), which is based on a prospective binary response model. Shown are the number of individuals required to detect a significant $G \times E$ interaction effect at $\alpha=0.01$ with a power of 80%. The horizontal solid line represents the sample size required for 80% power to detect a genetic main effect using a case-control design.

The interaction odds ratio was varied between 1.25 and 3, whereas the main effects of the genetic and environmental risk factors were 1.2 (a) and 1.5 (b). The disease model was defined by a recessive disease allele with frequency 0.3. The environmental risk factor had a prevalence of 30%. The baseline risk of the disease was 10%. The sample sizes to detect the genetic main effect, which were constant in the two scenarios, were 43,045, 16,196, and 19,860 in (a) for the cohort, case-control, and trio design and 7,712, 3,070 and 3,423 in (b), respectively. For a dominant disease allele, similar relations between required sample sizes are observed for the different designs (Dempfle et al. 2008. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

which give the required sample sizes for four different study designs (case–control, trio, case-only, and cohort) for varying effect sizes of the G×E interaction. Only for very weak marginal effects ($OR=1.2$, a) and at least moderate interactions ($OR>1.5$), the interaction is detectable with a smaller sample size than the marginal effect. But even for slightly larger marginal effects ($OR=1.5$, b) and weak to moderate interactions ($OR<2$), the sample size required to detect the interaction can be several fold higher than that required for detecting the marginal genetic effect. These examples are based on a level of significance (0.01) that might be used in a confirmatory study for testing one well-defined *a priori* hypothesis (e.g., one polymorphism within one gene). Sample sizes would be much higher for (exploratory) studies such as genome-wide association scans with hundreds of thousands of markers, as the correction for multiple testing requires much smaller levels of significance and thus much larger samples. In addition, these studies rely on linkage disequilibrium between the genotyped markers and potentially untyped disease alleles, and such indirect association studies may need much larger sample sizes (Dempfle et al. 2008).

Owing to limited number of samples, most sample sizes in genetics are small. This scenario describes very well the small number of cases in aggressive periodontitis association studies. The number of subjects in studies of chronic periodontitis tends to be larger, but variations do occur. The size of subjects clearly contributes to the differences in statistical power of the results, especially in a complex disease like periodontitis (Yoshie et al. 2007).

17.4.4 Case Selection

The principal issues with regard to case ascertainment revolve around the extent to which selection should be driven by manoeuvres that are designed to improve study power through enrichment for specific disease-predisposing alleles. These include efforts to minimize phenotypic heterogeneity or to focus on extreme and/or familial cases. Because the genetic architecture of most complex traits remains poorly understood, the

value of such efforts is hard to predict. In most circumstances, and particularly when the total GWA sample size has financial or operational constraints, efforts to enrich case selection are likely to improve power (McCarthy et al. 2008)

17.4.5 Choice of Controls

One consequence of the common-control design is the potential loss of power that is associated with the inability to exclude latent diagnoses of the phenotype of interest through intensive screening of controls. Fortunately, the consequences of misclassification bias are modest unless the trait is common, and any loss of power is recoverable by increasing the sample size (Fig. 17.7) (McCarthy et al. 2008).

In a case–control study, the manner in which the controls are ascertained (with respect to the phenotype of interest) has implications for the power of the study and for sample size. The panels on the left show estimates of power for a sample size of 2,000 cases and 2,000 controls and α (P value)=10^{−6}. Those on the right show the sample sizes (i.e., the number of case–control pairs) that are required for 80% power at the same threshold. In the upper panels, the disease of interest has a population prevalence of 5% (so that cases are ascertained purely from the top 5% of the population distribution); in the lower panels, the population prevalence is 20%. In each panel, power or sample size estimates are shown for a range of control selection thresholds, that is, the trait-distribution threshold that is used to define the controls. Under scenario A, controls are ascertained from the full distribution (i.e., population-based controls): A proportion (5% or 20%) will meet the criteria for being cases. Under scenario B, controls are ascertained only if they cannot be cases: They come from the residual part (bottom 80% or 95%) of the distribution. Under scenario C, hypernormal controls have been selected exclusively from the lowest 5% of the distribution. Each panel considers four potential susceptibility loci. Tracks in blue denote loci that account for 0.25% of overall trait variance; tracks in red denote loci that account for 1%.

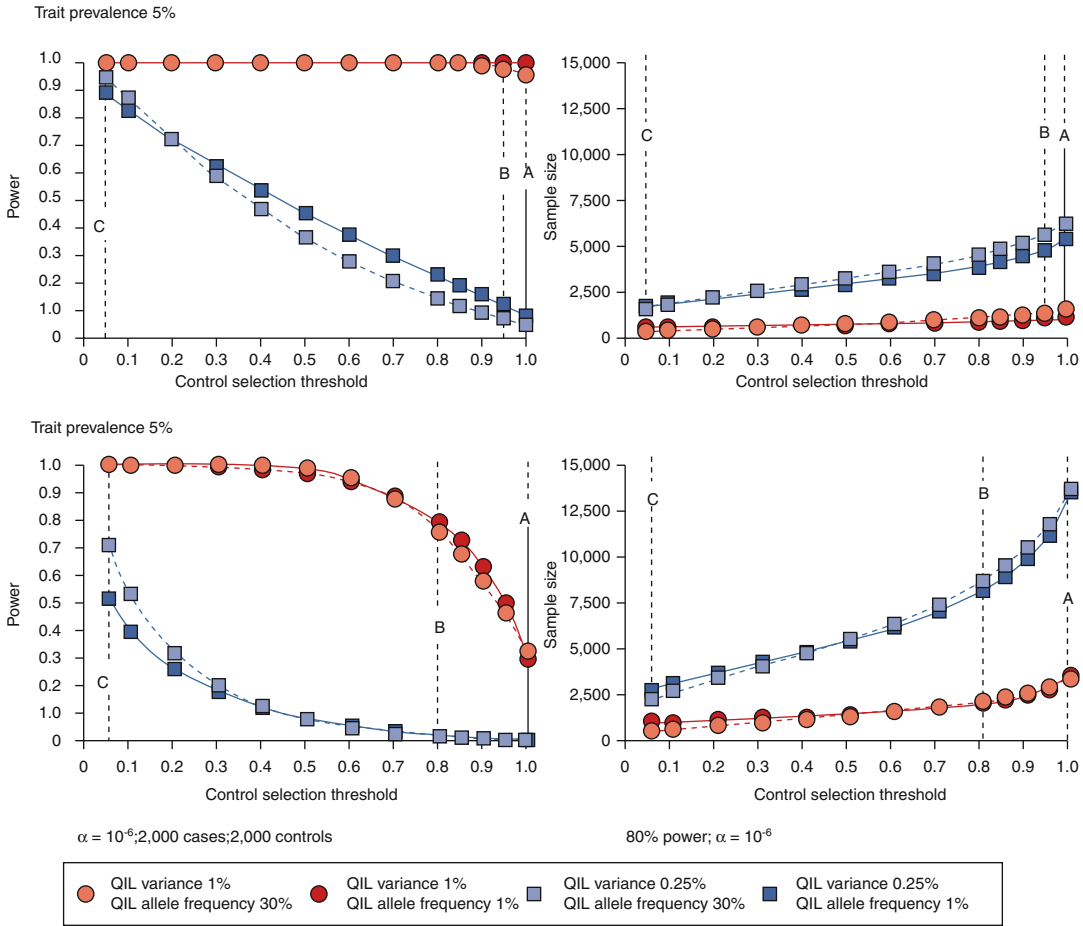


Fig. 17.7 The impact of selection by phenotype among controls on power and sample size

Light red and light blue symbols denote that the variant responsible is common (overall allele frequency 30%); red and dark blue symbols denote that the variant is rare (1%).

As expected, in all settings, scenario C is the most powerful strategy for given overall case–control sample size, and scenario A is the least powerful strategy. When the disease prevalence is modest (5%; upper panels), the distinctions between scenarios A and B are not large, and it will often be easier to increase sample size than to undertake detailed phenotypic examination of the controls to exclude latent cases. When the disease prevalence is higher (20%; lower panels), misclassification is more prevalent under scenario A, the adverse consequences of using population-based controls are more marked, and the advan-

tages of using hypernormal controls (scenario C), if available, are most obvious (McCarthy et al. 2008) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

Defining the appropriate controls for a case–control study in periodontitis still lacks clarity. Some reports generally described their control as healthy, while others specifically characterized controls as patients with gingivitis or slight periodontitis (Yoshie et al. 2007).

17.4.6 Capture of Complete Genetic Information

Apart from insufficient statistical power of most genetic association studies in periodontitis within

the last decade, most studies were unable to draw an unambiguous conclusion from their findings, even when the outcome was negative. Apart of the reasons mentioned above, this was also because most genetic association studies on periodontitis did not capture the complete genetic information of the region of interest. Only one or a few variants were genotyped in these studies, and usually, their selection had reasons that were more historical than biological. But a precise statement about a possible association of a genetic locus with a disease can only be made, if the full haplotype information has been assessed (Slatkin 2008). As findings of negative associations of single variants or haplotype blocks cannot rule out a potential disease association of the gene of interest, such studies have a very limited value. Therefore, as the outcome of an association study is always unclear, it is necessary that each genetic study needs to assess the haplotype information as completely as possible. Genotyping one variant simply provides no information on a possible disease association of the genetic locus when the outcome is negative (Schäfer et al. 2011).

17.4.7 Study Replication

Variants that are associated with complex diseases increase the disease risk rather modestly. With a sufficiently powered clinical analysis population, the hypothesis of the study can be generated. In other words, the explorative association study merely observes an association (or does not observe), thus creating the hypothesis, that a specific variant is/is not associated with the disease. The hypothesis obviously requires a test, which is the replication. This replication is of utmost importance and considered as the gold standard, which defines that each association is only as good as its replication. It goes without saying that the replication needs to be performed in an independent case-control population of the same phenotype (diagnosis criteria) and to be sampled from the same ethnic background. A repetition of the experiment with samples from different ethnic groups,

with different diagnosis criteria, or independent cases but the same controls is not a replication and does not test the hypothesis properly. Only after being confirmed can the hypothesis be validated in different sub-phenotypes or in different ethnic groups (Schäfer et al. 2011).

17.4.8 Data Presentation

Studies reporting DNA polymorphism-disease associations represent an important source of information on disease gene candidacy. Such studies are, however, extremely variable in terms of their design and of the statistical methodology employed. Guidelines are therefore proposed that are intended to promote the publication of scientifically meaningful disease association studies through the introduction of sensible methodological standards (Cooper et al. 2002).

Cooper et al. (2002) recommended that authors strive to ensure that papers describing putative associations between genotypes and phenotypes adhere as far as possible to the following guidelines:

- Ideally, even for an initial exploratory study, the heritability of the phenotype in question should already have been established. If this is not feasible, the assumption that genetic factors influence the phenotype in question must be properly justified.
- The choice of candidate gene(s) for study should be fully justified and firmly based upon a priori arguments. These arguments may either be positional or functional or both.
- In case-control studies, subjects should have been carefully matched by ethno-geographic origin and by any other potential confounding variables (e.g., social indicators) in order to avoid systematic differences in genetic composition between the two groups. The choice of matching criteria should be justified in terms of both the inclusion and (if applicable) the exclusion of confounders. If authors have attempted to correct for potential biases (e.g., through matching by genetic background), the

techniques employed must be properly described and/or referenced.

- If the study was not originally designed to confirm an existing biologically or clinically plausible hypothesis, the findings from the initial exploratory analysis should be replicated using an independent sample. Alternatively, validation could be achieved through functional studies involving *in vitro* or *in vivo* laboratory experimentation. Independent validation is particularly important in situations where an apparently significant association has been obtained only after post hoc stratification of data that would otherwise have yielded negative results. If it is not possible to obtain an independent sample or suitable experimental data, then the results of the exploratory analysis should be clearly presented and discussed as a novel hypothesis requiring further validation.
- Associations should preferably be reported in the form of effect size estimates and accuracy measures, such as odds ratios and their confidence limits. However, since *P* values are also commonly used in association studies to test against the null hypothesis of no association, authors are urged to report, and to reflect upon, the total number of associations that have been studied for the phenotype in question. The hypotheses being tested ought to be defined clearly so that a decision can be made as to whether correction for multiple testing is necessary or not
- Selective publication of positive findings inevitable leads to publication bias. Therefore, negative results of association studies will also be published under certain circumstances. In such situations, priority will be given to results that favour rejection of a previously published claim of association over negative findings on candidate genes for which no prior data existed. In any case, authors should provide power calculations in order to illustrate the effect sizes that have been excluded on the basis of their own negative data.
- The publication of an association identical or virtually identical (in terms of, e.g., population, clinical phenotype, and genotype) to one already

published and independently confirmed is justified only if some novel aspect is presented.

17.4.8.1 Reporting Genetic Association Studies: The STREGA Statement

The rapidly evolving evidence on genetic associations is crucial to integrating human genomics into the practice of medicine and public health. Genetic factors are likely to affect the occurrence of numerous common diseases, and therefore identifying and characterizing the associated risk (or protection) will be important in improving the understanding of aetiology and potentially for developing interventions based on genetic information (Little et al. 2009a, b).

The completion of sequencing of the human genome created high hopes for the translational potential of knowledge from genetic association studies. The combination of advances in genomic and related technologies and methods of conventional epidemiology has produced one of the most dynamic fields of medical research. Results of medical research need to be reported with sufficient accuracy and transparency to enable critical appraisal. Reporting guidelines aim to improve the research literature by the use of checklists with items deemed essential by multidisciplinary working groups. The STrengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement was developed to help authors to improve the reporting of cross-sectional, case-control, and cohort studies (von Elm et al. 2007). Genetic association studies that use one of these study designs present several specific challenges. As part of its road map for human genome epidemiology, the Human Genome Epidemiology Network set up a working group to provide guidance on how to report these studies. This initiative, called **STrengthening the REporting of Genetic Association (STREGA) studies**, elaborated on the STROBE Statement. The resulting STREGA Statement was recently published in several journals simultaneously. The STREGA checklist provides additions to 12 of the 22 items of the STROBE checklist to cover crucial aspects of genetic epidemiology reporting (Table 17.3) (von Elm et al. 2009).

Table 17.3 STREGA reporting recommendations, extended from STROBE Statement (Little et al. 2009a, b) (Reprinted with permission from John Wiley & Sons, Inc.)

Item	Item number	STROBE guideline	Extension for genetic association studies (STREGA)
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	
Introduction			
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any pre-specified hypotheses	State if the study is the first report of a genetic association, a replication effort, or both
Methods			
Study design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) Cohort study: give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. Case-control study: give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. Cross-sectional study: give the eligibility criteria and the sources and methods of selection of participants (b) Cohort study: for matched studies, give matching criteria and number of exposed and unexposed. Case-control study: for matched studies, give matching criteria and the number of controls per case	Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant
Variables	7	(a) Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	(b) Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be associated with population stratification (confounding by ethnic origin)
Data sources/measurement	8 ^a	(a) For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	(b) Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used and its version), error rates and call rates. State the laboratory/center where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches

Item	Item number	STROBE guideline	Extension for genetic association studies (STREGA)
Bias	9	(a) Describe any efforts to address potential sources of bias	(b) For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of the potential bias, and explain what approach was used to deal with this
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	If applicable, describe how effects of treatment were dealt with
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) Cohort study: if applicable, explain how loss to follow-up was addressed Case-control study: if applicable, explain how matching of cases and controls was addressed Cross-sectional study: if applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	State software version used and options (or settings) chosen (f) State whether Hardy–Weinberg equilibrium was considered and, if so, how (g) Describe any methods used for inferring genotypes or haplotypes (h) Describe any methods used to assess or address population stratification (i) Describe any methods used to address multiple comparisons or to control risk of false-positive findings (j) Describe any methods used to address and correct for relatedness among subjects
Results			
Participants	13 ^a	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful

(continued)

Table 17.3 (continued)

Item	Item number	STROBE guideline	Extension for genetic association studies (STREGA)
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

STREGA Strengthening the REporting of Genetic Association studies, *STROBE* Strengthening the Reporting of Observational Studies in Epidemiology

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies

Complementary information, on these areas, and the rationale for additional STREGA recommendations relating to selection of participants, choice of genes and variants selected, treatment effects in studying quantitative traits, statistical methods, relatedness, reporting of descriptive and outcome data, and issues of data volume are presented in **Table 17.4** (Little et al. 2009a, b).

The STREGA Statement has several strengths. First, it is based on existing guidance on reporting observational studies (STROBE). Second, it was developed from discussions of an interdisciplinary group that included epidemiologists, geneticists, statisticians, journal editors, and graduate students, thus reflecting a broad collaborative approach in terminology accessible to scientists from diverse disciplines. Finally, it explicitly describes the rationale for the decisions and has a clear plan for dissemination and evaluation. The STREGA recommendations are available at <http://www.strega-statement.org/> (Little et al. 2009a, b).

17.4.8.2 Strengthening the Reporting of Genetic Risk Prediction Studies: The GRIPS Statement

Genetic risk prediction studies typically concern the development and/or evaluation of models for the prediction of a particular health outcome, but there is considerable variation in their design, conduct, and analysis. Genetic risk models most frequently predict risk of disease, but they are also being investigated for the prediction of prognostic outcome, treatment response, or treatment side effects. Risk prediction models are used in

research and clinical settings to classify individuals into homogeneous groups, for example, for randomization in clinical trials and for targeting preventive or therapeutic interventions. The main study designs are cohort, cross-sectional, or case-control. The genetic risk factors often are single nucleotide polymorphisms, but other variants such as insertions/deletions, haplotypes, and copy number variations can be included as well. The risk models are based on genetic variants only or include both genetic and non-genetic risk factors. Risk prediction models are statistical algorithms, which can be simple genetic risk scores (e.g., risk allele counts), or be based on regression analyses (e.g., weighted risk scores or predicted risks) or on more complex analytic approaches such as support vector machine learning or classification trees. Papers on genetic risk prediction vary as to whether they present the development of a risk model only, the validation of one or more risk models only, or both development and validation of a risk model (Janssens et al. 2011a).

The GRIPS statement was developed by a multidisciplinary panel of 25 risk prediction researchers, epidemiologists, geneticists, methodologists, statisticians, and journal editors, seven of whom were also part of the STREGA initiative. This draft version was developed on the basis of existing reporting guidelines, namely, STREGA (Little et al. 2009a, b), REMARK (McShane et al. 2005), and STARD (Bossuyt et al. 2003). These were selected out of all available guidelines (see <http://www.equator-network.org>) because of their focus on observational study

Table 17.4 Rationale for inclusion of topics in the STREGA recommendations (Little et al. 2009a, b) and references therein (Reprinted with permission from John Wiley & Sons, Inc.)

Specific issue in genetic association studies	Rationale for inclusion in STREGA	Item(s) in STREGA	Specific suggestions for reporting
Main areas of special interest			
Genotyping errors (misclassification of exposure)	Non-differential genotyping errors will usually bias associations towards the null. When there are systematic differences in genotyping according to outcome status (differential error), bias in any direction may occur	8(b): Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory/centre where genotyping was performed. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches 13(a): Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful	Factors affecting the potential extent of misclassification (information bias) of genotype include the types and quality of samples, timing of collection and the method used for genotyping. When high throughput platforms are used, it is important to report not only the platform used but also the allele calling algorithm and its version. Different calling algorithms have different strengths and weaknesses. For example, some of the currently used algorithms are notably less accurate in assigning genotypes to single nucleotide polymorphisms with low minor allele frequencies (<0.10) than to single nucleotide polymorphisms with higher minor allele frequency. Algorithms are continually being improved. Reporting the allele calling algorithm and its version will help readers to interpret reported results, and it is critical for reproducing the results of the study given the same intermediate output files summarizing intensity of hybridization For some high throughput platforms, the user may choose to assign genotypes using all of the data from the study simultaneously, or in smaller batches, such as by plate. This choice can affect both the overall call rate and the robustness of the calls For case-control studies, whether genotyping was done blind to case-control status should be reported, along with the reason for this decision
Population stratification (confounding by ethnic origin)	When study sub-populations differ both in allele (or genotype) frequencies and disease risks, then confounding will occur, if these sub-populations are unevenly distributed across exposure groups (or between cases and controls)	12(h): Describe any methods used to assess or address population stratification	In view of the debate about the potential implications of population stratification for the validity of genetic association studies, transparent reporting of the methods used, or stating that none was used, to address this potential problem is important for allowing the empirical evidence to accrue Ethnicity information should be presented, as should genetic markers or other variables likely to be associated with population stratification. Details of case-family control designs should be provided, if they are used As several methods of adjusting for population stratification have been proposed, explicit documentation of the methods is needed

Modelling haplotype variation	In designs considered in this article, haplotypes have to be inferred because of lack of available family information. There are diverse methods for inferring haplotypes	12(g): Describe any methods used for inferring genotypes or haplotypes	When discrete 'windows' are used to summarize haplotypes, variation in the definition of these may complicate comparisons across studies, as results may be sensitive to choice of windows. Related 'imputation' strategies are also in use. It is important to give details on haplotype inference and, when possible, uncertainty. Additional considerations for reporting include the strategy for dealing with rare haplotypes, window size and construction (if used) and choice of software
Hardy–Weinberg equilibrium (HWE)	Departure from Hardy–Weinberg equilibrium may indicate errors or peculiarities in the data. Empirical assessments have found that 20–69% of genetic associations were reported with some indication about conformity with Hardy–Weinberg equilibrium, and that among some of these, there were limitations or errors in its assessment.	12(f): State whether Hardy–Weinberg equilibrium was considered and, if so, how	Any statistical tests or measures should be described, as should any procedure to allow for deviations from Hardy–Weinberg equilibrium in evaluating genetic associations.
Replication	Publications that present and synthesize data from several studies in a single report are becoming more common	3: State if the study is the first report of a genetic association, a replication effort, or both	The selected criteria for claiming successful replication should also be explicitly documented
Additional issues			
Selection of participants	Selection bias may occur if (1) genetic associations are investigated in one or more subsets of participants (sub-samples) from a particular study; or (2) there is differential non-participation in groups being compared; or, (3) there are differential genotyping call rates in groups being compared	6(a): Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant 13(a): Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful	Inclusion and exclusion criteria, sources and methods of selection of sub-samples should be specified, stating whether these were based on a priori or post-hoc considerations

(continued)

Table 17.4 (continued)

Specific issue in genetic association studies	Rationale for inclusion in STREGA	Item(s) in STREGA	Specific suggestions for reporting
Rationale for choice of genes and variants investigated	Without an explicit rationale, it is difficult to judge the potential for selective reporting of study results. There is strong empirical evidence from randomised controlled trials that reporting of trial outcomes is frequently incomplete and biased in favour of statistically significant findings. Some evidence is also available in pharmacogenetics.	7(b): Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be associated with population stratification (confounding by ethnic origin)	The scientific background and rationale for investigating the genes and variants should be reported For genome-wide association studies, it is important to specify what initial testing platforms were used and how gene variants are selected for further testing in subsequent stages. This may involve statistical considerations (for example, selection of <i>P</i> -value threshold), functional or other biological considerations, fine mapping choices, or other approaches that need to be specified Guidelines for human gene nomenclature have been published by the Human Gene Nomenclature Committee. Standard reference numbers for nucleotide sequence variations, largely but not only SNPs are provided in dbSNP, the National Center for Biotechnology Information's database of genetic variation. For variations not listed in dbSNP that can be described relative to a specified version, guidelines have been proposed.
Treatment effects in studies of quantitative traits	A study of a quantitative variable may be compromised when the trait is subjected to the effects of a treatment, for example, the study of a lipid-related trait for which several individuals are taking lipid-lowering medication. Without appropriate correction, this can lead to bias in estimating the effect and loss of power	9(b): For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of the potential bias, and explain what approach was used to deal with this 11: If applicable, describe how effects of treatment were dealt with	Several methods of adjusting for treatment effects have been proposed. As the approach to deal with treatment effects may have an important impact on both the power of the study and the interpretation of the results, explicit documentation of the selected strategy is needed

Statistical methods	Analysis methods should be transparent and replicable, and genetic association studies are often performed using specialized software	12(a): State software version used and options (or settings) chosen	12(b): State software version used and options (or settings) chosen
Relatedness	The methods of analysis used in family-based studies are different from those used in studies that are based on unrelated cases and controls. Moreover, even in the studies that are based on apparently unrelated cases and controls, some individuals may have some connection and may be (distant) relatives, and this is particularly common in small, isolated populations, for example, Iceland. This may need to be probed with appropriate methods and adjusted for in the analysis of the data	12(j): Describe any methods used to address and correct for relatedness among subjects	For the great majority of studies in which samples are drawn from large, non-isolated populations, relatedness is typically negligible and results would not be altered depending on whether relatedness is taken into account. This may not be the case in isolated populations or those with considerable inbreeding. If investigators have assessed for relatedness, they should state the method used and how the results are corrected for identified relatedness
Reporting of descriptive and outcome data	The synthesis of findings across studies depends on the availability of sufficiently detailed data	14(a): Consider giving information by genotype 15: Cohort study – report outcomes (phenotypes) for each genotype category over time Case-control study – report numbers in each genotype category Cross-sectional study – report outcomes (phenotypes) for each genotype category	

(continued)

Table 17.4 (continued)

Specific issue in genetic association studies	Rationale for inclusion in STREGA	Item(s) in STREGA	Specific suggestions for reporting
Volume of data	The key problem is of possible false-positive results and selective reporting of these. Type I errors are particularly relevant to the conduct of genome-wide association studies. A large search among hundreds of thousands of genetic variants can be expected by chance alone to find thousands of false positive results (odds ratios significantly different from 1.0)	12(i): Describe any methods used to address multiple comparisons or to control risk of false positive findings 16(d): Report results of any adjustments for multiple comparisons 17(b): If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken 17(c): If detailed results are available elsewhere, state how they can be accessed	Genome-wide association studies collect information on a very large number of genetic variants concomitantly. Initiatives to make the entire database transparent and available online may supply a definitive solution to the problem of selective reporting. Availability of raw data may help interested investigators reproduce the published analyses and also pursue additional analyses. A potential drawback of public data availability is that investigators using the data second-hand may not be aware of limitations or other problems that were originally encountered, unless these are also transparently reported. In this regard, collaboration of the data users with the original investigators may be beneficial. Issues of consent and confidentiality may also complicate what data can be shared, and how. It would be useful for published reports to specify not only what data can be accessed and where, but also briefly mention the procedure. For articles that have used publicly available data, it would be useful to clarify whether the original investigators were also involved and if so, how The volume of data analysed should also be considered in the interpretation of findings Examples of methods of summarizing results include giving distribution of <i>P</i> -values (frequentist statistics), distribution of effect sizes and specifying false discovery rates

STREGA STrengthening the Reporting of Genetic Association studies, *SNP* single nucleotide polymorphism

designs and genetic factors (STREGA), prediction models (REMARK), and test evaluation (REMARK and STARD). The final version of the checklist is presented in **Table 17.5**.

The GRIPS guidelines were developed to improve the transparency, quality, and completeness of the reporting of genetic risk prediction studies. GRIPS does not prescribe how studies should be designed, conducted, and analysed, and therefore the guidelines should not be used to assess the quality of empirical studies. The guidelines should only be used to check whether all essential items are adequately reported (Janssens et al. 2011a).

17.5 Limitations of Genetic Association Studies

Genome-wide association (GWA) studies are the most widely used contemporary approach to relate genetic variation to phenotypic diversity. Over the past 2 years, these studies have identified statistical association between hundreds of loci across the genome and common complex traits. The results of these studies have substantially increased our understanding of the diverse molecular pathways underlying specific human diseases. However, GWA studies have several limitations. First, there is great difficulty moving beyond mere statistical associations to identifying the functional basis of the link between a genomic interval and a given complex trait. Second, SNP associations identified in one population frequently are not transferable to members of other populations. Third, the bulk of the heritable fraction of complex traits has not been accounted for in recent GWA studies. This last point is probably explained by the fact that GWA studies do not capture information about rare variants and have limited statistical power to detect small gene–gene and gene–environment interactions (Frazer et al. 2009).

Understanding the heritability of genetic diseases requires a more comprehensive assessment of human genetic variation. Human genomes are rich in structural diversity, but the discovery and genotyping of this type of variation has lagged

far behind those of the SNP. Although there has been a tremendous push to close this gap over the past 4 years, two aspects remain understudied:

- The first is the exploration of the landscape and impact of large variants (deletions, duplications, and inversions) that are individually rare but collectively common in the human population. The available data suggest that these variants are under strong selection, affect transcription, and contribute to a variety of different diseases. These genomic imbalances represent a special class of rare variants that can potentially affect many genes and pathways in a single individual. Not only are large numbers of cases and controls required to assess the clinical significance of particular events, but the modelling of other forms of genetic variation in this sensitized background of localized haploidy or triploidy remains largely unexplored.
- The second aspect involves the several hundred genes that map to regions of copy number polymorphic (CNP) duplications. Available data suggest that these genes are highly variable among individuals, and because of their repetitive and multicopy nature are considered inaccessible by most existing genotyping and sequencing technologies. There is a pressing need to characterize not only copy number but also the sequence content and structural arrangement in these diverse regions of our genome. Such regions are more likely to be subject to recurrent mutations and be inadequately assayed by a correlated neighbouring SNP. It is therefore premature to conclude that CNPs have limited impact in terms of common disease until these more complex regions are tested (Eichler et al. 2010)

Only if the major contributory elements are fully understood, which jointly predisposes people to the specific individual disease phenotype, a risk prediction or an individualized therapy will be possible. Here, to meet with the current discussion on the best future research strategies to elucidate the missing heritability would be inadequate, and we would like to refer to the cited literature. Detecting associations with common genetic variants is still considered as the first and also the most

Table 17.5 Reporting recommendations for evaluations of risk prediction models that include genetic variants (Janssens et al. 2011b) (Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

Title and abstract		
	1	(a) Identify the article as a study of risk prediction using genetic factors. (b) Use recommended keywords in the abstract: genetic or genomic, risk, prediction
Introduction		
Background and rationale	2	Explain the scientific background and rationale for the prediction study
Objectives	3	Specify the study objectives and state the specific model(s) that is/are investigated. State if the study concerns the development of the model(s), a validation effort or both
Methods		
Study design and setting	4 ^a	Specify the key elements of the study design and describe the setting, locations and relevant dates, including periods of recruitment, follow-up and data collection
Participants	5 ^a	Describe eligibility criteria for participants, and sources and methods of selection of participants
Variables: definition	6 ^a	Clearly define all participant characteristics, risk factors and outcomes. Clearly define genetic variants using a widely used nomenclature system
Variables: assessment	7 ^a	(a) Describe sources of data and details of methods of assessment (measurement) for each variable. (b) Give a detailed description of genotyping and other laboratory methods
Variables: coding	8	(a) Describe how genetic variants were handled in the analyses. (b) Explain how other quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Analysis: risk model construction	9	Specify the procedure and data used for the derivation of the risk model. Specify which candidate variables were initially examined or considered for inclusion in models. Include details of any variable selection procedures and other model-building issues. Specify the horizon of risk prediction (e.g. 5-year risk)
Analysis: validation	10	Specify the procedure and data used for the validation of the risk model
Analysis: missing data	11	Specify how missing data were handled
Analysis: statistical methods	12	Specify all measures used for the evaluation of the risk model including, but not limited to, measures of model fit and predictive ability
Analysis: other	13	Describe all subgroups, interactions and exploratory analyses that were examined
Results		
Participants	14 ^a	Report the numbers of individuals at each stage of the study. Give reasons for nonparticipation at each stage. Report the number of participants not genotyped, and reasons why they were not genotyped
Descriptives: population	15 ^a	Report demographic and clinical characteristics of the study population, including risk factors used in the risk modelling
Descriptives: model estimates	16	Report unadjusted associations between the variables in the risk model(s) and the outcome. Report adjusted estimates and their precision from the full risk model(s) for each variable
Risk distributions	17 ^a	Report distributions of predicted risks and/or risk scores
Assessment	18	Report measures of model fit and predictive ability, and any other performance measures, if pertinent
Validation	19	Report any validation of the risk model(s)
Other analyses	20	Present results of any subgroup, interaction or exploratory analyses, whenever pertinent

(continued)

Table 17.5 (continued)

Discussion		
Limitations	21	Discuss limitations and assumptions of the study, particularly those concerning study design, selection of participants, measurements and analyses, and discuss their impact on the results of the study
Interpretation	22	Give an overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies and other relevant evidence
Generalizability	23	Discuss the generalizability and, if pertinent, the healthcare relevance of the study results
Other		
Supplementary information	24	State whether databases for the analyzed data, risk models and/or protocols are or will become publicly available, and if so, how they can be accessed
Funding	25	Give the source of funding and the role of the funders for the present study. State whether there are any conflicts of interest

^aMarked items should be reported for every population in the study

straightforward and cost-efficient step, towards a better understanding of the genetic architecture of a complex disease (Schäfer et al. 2011).

References

- Armitage GC. Development of a classification system for periodontal diseases and conditions. *Ann Periodontol*. 1999;1:1–6.
- Barros SP, Offenbacher S. Epigenetics: connecting environment and genotype to phenotype and disease. *J Dent Res*. 2009;5:400–8.
- Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen TS, Thomson JA. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol*. 2010;28:1045–8.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326:41–4.
- Cooper DN, Nussbaum RL, Krawczak M. Proposed guidelines for papers describing DNA polymorphism-disease associations. *Hum Genet*. 2002;110(3):207–8.
- Craig J. Complex diseases: Research and applications. *Nature Education*, 2008;1(1). Available at: <http://www.nature.com/scitable/topicpage/complex-diseases-research-and-applications-748>.
- Dempfle A, Scherag A, Hein R, Beckmann L, Chang-Claude J, Schäfer H. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. *Eur J Hum Genet*. 2008;16:1164–72.
- Dibbens LM, Heron SE, Mulley JC. A polygenic heterogeneity model for common epilepsies with complex genetics. *Genes Brain Behav*. 2007;6:593–7.
- Dumitrescu AL, Kobayashi J. Genetic variants in periodontal health and disease. 1st ed. Berlin/Heidelberg: Springer Germany; 2009.
- Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet*. 2010;11:446–50.
- Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet*. 2009;10:241–51.
- Gomez RS, Dutra WO, Moreira PR. Epigenetics and periodontal disease: future perspectives. *Inflamm Res*. 2009;58:625–9.
- Hirst M, Marra MA. Epigenetics and human disease. *Int J Biochem Cell Biol*. 2009;1:136–46.
- Hodge SE. Linkage analysis versus association analysis: distinguishing between two models that explain disease-marker associations. *Am J Hum Genet*. 1993;53:367–84.
- Hodge DR, Peng B, Cherry JC, Hurt EM, Fox SD, Kelley JA, et al. Interleukin 6 supports the maintenance of p53 tumor suppressor gene promoter methylation. *Cancer Res*. 2005;65:4673–82.
- Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet*. 2005;6:287–98.
- Iacopino AM. Epigenetics: new explanations for old problems? *J Can Dent Assoc*. 2010;76:a76.
- Imai K, Ochiai K. Role of histone modification on transcriptional regulation and HIV-1 gene expression: possible mechanisms of periodontal diseases in AIDS progression. *J Oral Sci*. 2011;53:1–13.
- Janssens AC, Ioannidis JP, van Duijn CM, Little J, Khoury MJ, GRIPS Group. Strengthening the reporting of genetic risk prediction studies: the GRIPS statement. *Genet Med*. 2011a;13:453–6.
- Janssens AC, Ioannidis JP, van Duijn CM, Little J, Khoury MJ, Human Genome Epidemiology Network. Strengthening the reporting of genetic risk prediction studies: the GRIPS statement. *Eur J Hum Genet*. 2011b;19:833–6.

- Jia LY, Tam PO, Chiang SW, Ding N, Chen LJ, Yam GH, Pang CP, Wang NL. Multiple gene polymorphisms analysis revealed a different profile of genetic polymorphisms of primary open-angle glaucoma in northern Chinese. *Mol Vis*. 2009;15:89–98.
- Kinane DF, Hart TC. Genes and gene polymorphisms associated with periodontal disease. *Crit Rev Oral Biol Med*. 2003;14:430–49.
- Laine ML, Crielgaard W, Loos BG. Genetic susceptibility to periodontitis. *Periodontol* 2000. 2012;58:37–68.
- Lee JH, Moore JH, Park SW, Jang AS, Uh ST, Kim YH, Park CS, Park BL, Shin HD. Genetic interactions model among Eotaxin gene polymorphisms in asthma. *J Hum Genet*. 2008;53:867–75.
- Listgarten J, Damaraju S, Poulin B, Cook L, Dufour J, Driga A, Mackey J, Wishart D, Greiner R, Zanke B. Predictive models for breast cancer susceptibility from multiple single nucleotide polymorphisms. *Clin Cancer Res*. 2004;10:2725–37.
- Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N. STrengthening the REporting of Genetic Association studies (STREGA) – an extension of the STROBE statement. *Eur J Clin Invest*. 2009a;39:247–66.
- Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N, STrengthening the REporting of Genetic Association Studies. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. *PLoS Med*. 2009b;6:e22.
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9:356–69.
- McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, et al. Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol*. 2005;23:9067–72.
- Moreira PR, Lima PMA, Sathler KOB, Imanishi SA, Costa JE, Gomes RS, et al. Interleukin-6 expression and gene polymorphism are associated with severity of periodontal disease in a sample of Brazilian individuals. *Clin Exp Immunol*. 2007;148:119–26.
- Nakao M. Epigenetics: interaction of DNA methylation and chromatin. *Gene*. 2001;278:25–31.
- Nikolopoulos GK, Dimou NL, Hamodrakas SJ, Bagos PG. Cytokine gene polymorphisms in periodontal disease: a meta-analysis of 53 studies including 4178 cases and 4590 controls. *J Clin Periodontol*. 2008;35:754–67.
- Nose M. A polygene network model for the complex pathological phenotypes of collagen disease. *Pathol Int*. 2011;61:619–29.
- Offenbacher S, Barros SP, Beck JD. Rethinking periodontal inflammation. *J Periodontol*. 2008;79:1577–84.
- Pers TH, Hansen NT, Lage K, Koefoed P, Dworzynski P, Miller ML, Flint TJ, Møllerup E, Dam H, Andreassen OA, Djurovic S, Mølle I, Børghlum AD, Werge T, Purcell S, Ferreira MA, Kouskoumvekaki I, Workman CT, Hansen T, Mors O, Brunak S. Meta-analysis of heterogeneous data sources for genome-scale identification of risk genes in complex phenotypes. *Genet Epidemiol*. 2011;35:318–32.
- Schäfer AS, Jepsen S, Loos BG. Periodontal genetics: a decade of genetic association studies mandates better study designs. *J Clin Periodontol*. 2011;38:103–7.
- Shen CD, Zhang WL, Sun K, Wang YB, Zhen YS, Hui RT. Interaction of genetic risk factors confers higher risk for thrombotic stroke in male Chinese: a multicenter case-control study. *Ann Hum Genet*. 2007;71:620–9.
- Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*. 2008 Jun;9(6):477–85.
- Stabholz A, Soskolne WA, Shapira L. Genetic and environmental risk factors for chronic periodontitis and aggressive periodontitis. *Periodontol* 2000. 2010;53:138–53.
- Stenvinkel P, Karimi M, Johansson S, Axelsson J, Suliman M, Lindholm B, et al. Impact of inflammation on epigenetic DNA methylation: a novel risk factor for cardiovascular disease? *J Intern Med*. 2007;261:488–99.
- Thakkestian A, McElduff P, D'Este C, Duffy D, Attia J. A method for meta-analysis of molecular association studies. *Stat Med*. 2005;24:1291–306.
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453–7.
- von Elm E, Moher D, Little J, STREGA collaboration. Reporting genetic association studies: the STREGA statement. *Lancet*. 2009;374(9684):98–100.
- Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, Rayner NW, Shields B, Owen KR, Hattersley AT, Frayling TM. Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLoS Med*. 2006;3:e374.
- Xiao R, Boehnke M. Quantifying and correcting for the winner's curse in genetic association studies. *Genet Epidemiol*. 2009;33:453–62.
- Xiao R, Boehnke M. Quantifying and correcting for the winner's curse in quantitative-trait association studies. *Genet Epidemiol*. 2011;35:133–8.
- Yoshie H, Galicia JC, Kobayashi T, Tai H. Genetic polymorphisms and periodontitis. *Int Congr Ser*. 2005;1284:131–9.
- Yoshie H, Kobayashi T, Tai H, Galicia JC. The role of genetic polymorphisms in periodontitis. *Periodontol* 2000. 2007;43:102–32.
- Zhang J, Sun X, Xiao L, Xie C, Xuan D, Luo G. Gene polymorphisms and periodontitis. *Periodontol* 2000. 2011;56:102–24.

In order to assess prevalence and severity of gingival and/or periodontal lesions, a variety of index systems have been developed. Index systems were first applied in epidemiologic studies. However, the application of some systems became more and more accepted in clinical practice as a part of the individual patient's documentation (Lang 1998).

Some guidelines for indices evaluation were defined by Hazen (1974):

1. An index must be simple to use and permit the study of large number of persons with a minimum of time and cost.
2. Criteria defining the components of the index should be clear and understandable to promote accuracy and reproducibility.
3. A severity index should be equally sensitive throughout and should indicate in a meaningful way the clinical stages of the disease process.
4. The index should be amenable to statistical analysis.

Clinical indices represent methods for converting the observed severity of parameters of interest into a numerical form, which is amenable to standard analysis. In order for an index to be useful, there must be an established relationship between severities as defined by the index criteria and actual clinical changes accompanying the progression or regression of the disease (Barnett 1996).

Indices may be classified into five categories: (1) indices for the inflammatory changes of the gingiva; (2) indices for the loss of periodontal

tissues; (3) oral hygiene indices; (4) miscellaneous indices; and (5) treatment needs indices (Lang 1998).

18.1 Evaluation of Inflammatory Changes of the Gingiva

Clinical indices for evaluating gingivitis should meet four basic criteria. The indices should be simple to use with a minimum of time and expense, the criteria for assessing the state of gingivitis should be clear and understandable, the clinical manifestations of the disease process should be measurable at all stages, and the data from the index should be amenable to statistical treatment. Accurate indices to assess the severity of gingivitis and the localized effect of treatment must be used. It has been suggested that those indices which rely upon subjective examiner observation of the colour of tissue or an estimate of the degree of tissue swelling can be supplemented by methods which tend to use more objective measures such as bleeding upon probing or the study of gingival crevicular fluid (Fischman 1988).

All gingival indices have relied on one or more of the following criteria: gingival colour, gingival contour, gingival bleeding, extent of gingival involvement, and gingival crevicular flow (Fischman 1988). Also, most indices have assigned a number to each of the criteria

evaluated. Although numerical indices are the easiest to use in clinical trials and epidemiological surveys, they are limited in that they do not constitute a ratio scale, that is, a two is not necessarily twice as much inflammation as a one. Another consideration with gingival indices is that they only evaluate the condition of the soft tissue and may or may not be directly or proportionally related to underlying periodontal destruction (Ciancio 1986).

However, marked differences in the description of the evaluation of the gingival inflammation exist (Hazen 1974). For example, Mühlemann and Son (1971) use bleeding upon gentle probing as a criterion for slight inflammation, and the sulcus bleeding index has been devised on this premise. Silness and Loe (1963) use early colour changes of the tissue and no bleeding on probing as their criteria for these early changes (Hazen 1974).

Primarily for use in clinical studies, several indices or grading systems have been developed to formally characterized the appearance of bleeding on gentle manipulation of inflamed gingiva. Since the amount of pressure used during probe insertion can influence whether or not a site bleeds, it may be advisable to use controlled-force pressure-sensitive periodontal probes when bleeding indices are taken (**Fig. 18.1**) (Armitage et al. 1994; Müller and Heinecke 2002).

There is limited and contradictory information regarding the effect of time on gingival bleeding after repeat probing (Birkeland and Jorkjend 1975; Janssen et al. 1986; Van der

Velden 1980). Müller and Barrieshi-Nusair (2005) assessed agreement and association of gingival bleeding after repeated probing at different time intervals in subjects with gingivitis. Systematic periodontal probing to the bottom of the 'sulcus' was conducted at six sites of every tooth present with a pressure-controlled probe (ClickProbe 1395, KerrHawe, Bioggio, Switzerland). The probe has a tip diameter of 0.5 mm, and probing force is, according to the manufacturer, 0.25 N when the probe bends with a palpable click. Thus, pressure was about 1.27 MPa. No time limit for assessing presence or absence of BOP was preset because bleeding may be delayed in certain individuals (Nowicki et al. 1981). Probing was repeated in different quadrants immediately (following thorough mouth rinsing with tap water) after the first probing, after 1, 4, and 24 h. Four sampling schemes were randomly assigned to volunteers. For a given quadrant probed immediately after the first probing, the contralateral quadrant in the opposite jaw was probed after 1 h. Probing was repeated after 4 h in the quadrant of the same side in the opposite jaw, and in the remaining quadrant, after 24 h. It was concluded that a certain degree of agreement of site-specific bleeding scores in subjects with plaque-induced gingivitis could be observed only if probing was repeated at once. Adjusted associations between repeat BOP were weak in general but strongest immediately after first probing. There appears to be a significant effect of probing itself, which

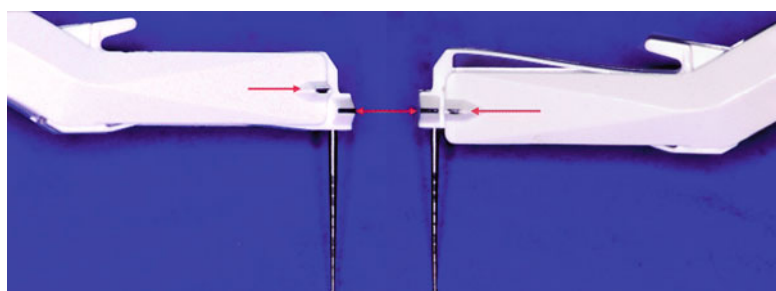


Fig. 18.1 *Left:* TPS probe without pressure applied. *Right:* TPS probe with a 0.2 N probing force applied, i.e., when the two black marks are at level with each other. The TPS probe (Vivadent Ets., Schaan, Liechtenstein) has a

diameter of 0.4 mm and is calibrated for a 0.2 N probing force (Abrahamsson and Soldini 2006. Reprinted with permission from John Wiley & Sons, Inc.)

may last for more than 1 h, whereas 24-h results are obtained under different conditions (Müller and Barrieshi-Nusair 2005). These observations may corroborate the findings of investigators who reported increased bleeding frequencies for up to 2 h after initial probing in subjects with gingivitis (Birkeland and Jorkjend 1975; Van der Weijden et al. 1994b), but may be in contrast to others who examined nonsurgically treated periodontitis patients with increased probing depths, and did not observe an increased percentage of bleeding sites when probing was repeated after 15 (Van der Velden 1980) or 100 min (Janssen et al. 1986). However, investigators employed different probes, probing forces, probing techniques, and frequently used inappropriate statistics (Müller and Barrieshi-Nusair 2005).

Gingival bleeding is an objective sign of inflammation in the gingival connective tissue. Bleeding occurs because of frequent micro-ulcerations in the epithelium that lines the soft-tissue wall of a periodontal pocket. Gingival bleeding is not a diagnosis; it does not distinguish between different forms of periodontal diseases and is not pathognomonic of any one form of the disease. Indeed, gingival bleeding is associated with various forms of periodontal diseases including gingivitis, acute necrotizing ulcerative gingivitis, juvenile periodontitis, adult periodontitis, rapidly progressive periodontitis, and refractory periodontitis (Newbrun 1996). However, it has been suggested that bleeding on probing may be an earlier and more reliable sign than visual indices for detection of a deviation from health (Ciancio 1986).

The choice of index should depend on the purpose of a study. For epidemiological surveys, partial recording of selected teeth or sites may be sufficient. On the other hand, for research and clinical trials, a quantitative measurement of bleeding is more informative than a dichotomous index of presence or absence of bleeding on stimulation. For patient education and motivation, a dichotomous index will suffice. Although some clinical investigators may favour a given index to the exclusion of all others, a variety of indices could be appropriate for use in clinical trials (Bessa Rebelo et al. 2011).



Fig. 18.2 Moderate gingivitis. Heavy bleeding can be observed as the papillary bleeding index is recorded

18.1.1 Papillary Marginal Attached (PMA) Index (Massler 1967)

The PMA index, developed by Schour and Massler (1948) and described by Massler (1967), was intended to survey gingival changes in relatively large groups for epidemiologic purposes and study. The basic philosophy used in the development of PMA index was very similar to the DMF index, that is, the number of gingival units affected was counted rather than the severity of the inflammation. However, the severity of the inflammation is also included in the records (Fig. 18.2) (Massler 1967).

The degree of gingivitis for each of the papillary (P), marginal (M), and attached (A) gingival units was defined as follows:

P: 0 normal

- 1+** mild papillary engorgement
- 2+** obvious increase in size of gingival papilla, haemorrhage on pressure
- 3+** excessive increase in size with spontaneous haemorrhage
- 4+** necrotic papilla
- 5+** atrophy and loss of papilla (through inflammation)

M: 0 normal

- 1+** engorgement, slight increase in size, no bleeding
- 2+** obvious engorgement, bleeding upon pressure
- 3+** swollen collar, spontaneous haemorrhage, beginning infiltration into attached gingivae
- 4+** necrotic gingivitis

5+ recession of the free marginal gingivae below the CEJ due to inflammatory changes

A: 0 normal, pale rose, stippled

1+ slight engorgement with loss of stippling; change in colour may or may not be present

2+ obvious engorgement of attached gingivae with marked increase in redness, pocket formation present.

3+ advanced periodontitis, deep pockets evident

The number of affected papillary units, marginal units and attached units were counted for each patient and recorded as follows: P-M-A=10-5-0 (15). Preference has been given to recording these numbers separately rather as a sum since the same sum can have different meanings: P-M-A=5-5-5 (15) (Massler 1967). Its major purpose has been the evaluation of gingival inflammation in children (Bessa Rebelo et al. 2011).

18.1.2 Papillary Marginal Gingivitis Index (PMGI)

This is a combination of the PMA index of Massler and Schour and the gingival index (GI) by Löe and Silness. Only the papillae and margins on the facial and lingual gingival of natural teeth are scored on a scale of 0–3, as follows:

Score 0=no inflammation or normal gingival

Score 1=mild inflammation or slight change in colour (erythema) and little change in structure

Score 2=moderate inflammation with moderate glazing, redness, oedema, and enlargement, as well as bleeding on pressure with a blunt instrument

Score 3=severe inflammation with marked redness and enlargement as well as tendency for spontaneous bleeding and ulceration

The individual score is the number of all inflammatory scores divided by the number of papillary and marginal units examined per subjects.

18.1.3 Sulcus Bleeding Index (SBI) (Mühlemann and Son 1971)

This is proposed by Mühlemann and Son (1971) based on the assumption that bleeding from the

sulcus was the earliest clinical symptom of gingivitis and that it even precedes discoloration and swelling (Lang 1998). The major characteristics of the SBI were:

1. For all severity scores, diagnosis of inflammation is made only if bleeding occurs upon gentle probing of sulcus.
2. Apparently healthy gingival units without colour changes or swelling are diagnosed as inflamed—score 1—if bleeding occurs upon gentle probing.
3. The SBI uses 8 maxillary and mandibular anterior teeth, and 4 gingival units are scored for each tooth: M labial, M lingual, P mesial, P distal (Lang 1998).

The following degrees of inflammation were differentiated:

Score 0=normal-appearing gingival, no bleeding upon probing

Score 1=no colour or contour changes but bleeding upon probing

Score 2=bleeding on probing, colour change (reddening), and no oedematous contour changes

Score 3=bleeding on probing, colour change, mild inflammatory oedema

Score 4=bleeding on probing, colour change, severe inflammatory oedema

Score 5=spontaneous bleeding on probing, colour change, very severe inflammatory oedema with or without ulceration (Engelberger et al. 1983)

For calculations, Lavigne (1998) suggested that it is necessary to probe the 4 gingival units mentioned above, wait 30 s after probing before scoring, total scores per tooth and divide by 4, calculate SBI for individual: total individual tooth scores, and divide by the number of teeth.

Although the gingival index (Loe and Silness 1963) and sulcus bleeding index (Mühlemann and Son 1971) have been widely used as indicators of periodontal status, there is some disagreement among investigators as to their meaning and significance. A clinical study was undertaken to monitor the occurrence of gingival bleeding, oedema, and change in colour in subjects with and without periodontal disease, and it was found that the combinations of these clinical symptoms often did not correspond exactly with an index score. It is therefore suggested that any study of periodontal

disease should be based on fundamental criteria, such as bleeding or oedema, rather than on composite indices (Benamghar et al. 1982).

18.1.4 Papillary Bleeding Index (PBI) (Saxer and Mühlemann 1975)

The papillary bleeding index was first introduced by Saxer and Mühlemann (1975), as cited by Mühlemann (1977). This index permits both immediate evaluation of the patient's gingival condition and his motivation, based upon the actual bleeding tendency of the gingival papillae (Bessa Rebelo et al. 2011). The rationale for this index is that marginal periodontitis and alveolar bone loss begin interproximally, and the effectiveness of preventive procedures is related to the presence or absence of interdental plaque. The results are translated into numerical scores which are easily comprehensible by the patient (Lang 1998). A periodontal probe is inserted into the gingival sulcus at the base of the papilla on the mesial aspect and then moved coronally to the papilla tip (Bessa Rebelo et al. 2011). This is repeated on the distal aspect of the papilla. The following scores are noted:

Score 0 = no bleeding

Score 1 = only one bleeding point appearing

Score 2 = several isolated bleeding points or a small blood area appearing

Score 3 = interdental triangle filled with blood soon after probing

Score 4 = profuse bleeding when probing, blood spreads towards the marginal papilla

Engelberger et al. (1983) showed that the papilla bleeding index (PBI) was positively highly statistically significant correlated with all of the other clinical indices sulcus bleeding index scores and gingival crevicular fluid levels. In contrast, between PBI and probing depth, and between SBI and probing depth, only weak correlations were found. The comparison of the PBI with the histological determination of inflammation revealed a clear increase in the absolute amount of inflammatory infiltrate as PBI scores increased. In addition, there was a definitive shift in the intensity of the infiltration; moderate and severe areas of infiltration became more common as PBI scores went up.

18.1.5 Modified Papillary Bleeding Index (MPBI)

Barnett et al. (1980) modified the PBI index (Mühlemann 1977) by stipulating that the periodontal probe should be gently placed in the gingival sulcus at the mesial line angle of the tooth surface to be examined and carefully swept forward into the mesial papilla (Bessa Rebelo et al. 2011). They timed the appearance of bleeding and graded it as follows:

0 = no bleeding within 30 s of probing

1 = bleeding between 3 and 30 s of probing

2 = bleeding within 2 s of probing

3 = bleeding immediately upon probe placement

The mesial papillae of all teeth present from the second molar to the lateral incisor were assessed. Indices were derived for the maxillary left and mandibular right buccal segments and the maxillary right and mandibular left lingual segments, and from these, a full-mouth index was calculated. This distribution of test sites was utilized since each mesial papilla could only be tested once, that is, from either the buccal or lingual side. They showed that the modified PBI may be more sensitive than the visual aspects of the GI in assessing changes in gingival health (Bessa Rebelo et al. 2011).

18.1.6 Papillary Bleeding Score (PBS)

This is performed using a Stim-U-dent®, which is inserted interproximally (Loesche 1979). Essentially, the PBS expands the score 2 of the gingival index (Silness and Løe 1963) into three recognized clinical conditions. The criteria are:

0 = healthy gingiva, no bleeding upon insertion of Stim-U-dent® interproximally

1 = oedematous, reddened gingiva, no bleeding upon insertion of Stim-U-Dent® interproximally

2 = bleeding, without flow, upon insertion of Stim-U-dent® interproximally

3 = bleeding, with flow, along gingival margin upon insertion of Stim-U-dent® interproximally

4 = copious bleeding upon insertion of Stim-U-dent® interproximally

5 = severe inflammation, marked redness and oedema, tendency to spontaneous bleeding

The PBS is determined on all papillae anterior to the second molars.

18.1.7 Modified Papillary Bleeding Score (PBS)

Modified version of the papillary bleeding score described by Loesche (1979) was used by Clements et al. (2003) who explored the treatment effects of material stiffness and frequency of appliance change when using clear, sequential, removable appliances (aligners). Stim-U-dent® plaque removers were inserted into the facial aspect of each interproximal site, and a score from 1 to 5 was assigned: 1, no bleeding; 2, slight bleeding; 3, bleeding with flow; 4, intermediate bleeding (copious); 5, severe inflammation (spontaneous bleeding). A score was recorded for each interdental papilla, and these scores were averaged to obtain a papillary bleeding score for each patient.

18.1.8 Gingival Index (Silness and Löe 1963)

The patient's gingival condition can also be scored according to the criteria of a gingival index (GI) system proposed by Silness and Löe (1963). This index is based on the clinical characteristics of different grades of gingival inflammation. The bleeding is assessed by probing gently along the wall of soft tissue of the gingival sulcus (Bessa Rebelo et al. 2011). The criteria for the gingival index system are:

Score 0 = absence of inflammation

Score 1 = mild inflammation: slight change in colour and little change in texture, no bleeding

Score 2 = moderate inflammation: moderate glazing, redness, oedema and hypertrophy, bleeding on pressure

Score 3 = severe inflammation: marked redness and hypertrophy, tendency to spontaneous bleeding, ulceration

Gingiva at six teeth representing the six segments of the jaws is examined: 16, 12, 24, 36, 32, and 44.

In order to circumvent the problems of mixing quality and extension of the diseased, Silness and

Löe (1963) gingival index refers to the individual tooth surface as the unit area, and, consequently, the criteria for the different scores have been made strictly qualitative. Each gingival unit (buccal, lingual, mesial, and distal) of the individual tooth is given a score from 0 to 3, called the GI for the area. The scores from the four areas of the tooth are added and divided by four to give the GI for the tooth. The scores of the individual teeth (incisors, premolars, and molars) may be grouped to designate the GI for the group of teeth. Finally, by adding the indices for the teeth and dividing by the number of teeth examined, the GI for the patient is obtained. The index for the patient is thus an average score for the areas examined (Löe 1967). A score of 0.1–1.0 = mild inflammation, 1.1–2.0 = moderate inflammation, and 2.1–3.0 signifies severe inflammation. The GI has been used frequently in clinical trials of therapeutic agents (Bessa Rebelo et al. 2011).

Originally, separate scores should be given for four aspects of each tooth (buccal, lingual, mesial and distal). Later, however, investigators have often given separate scores for six aspects of each tooth (mesio-buccal, mid-buccal, disto-buccal, mesio-lingual, mid-lingual, disto-lingual) (Egelberg 1999).

18.1.9 Modified Gingival Index (MGI)

The modified gingival index (MGI), devised by Lobene et al. (1986), introduced changes in the criteria of the gingival index (Silness and Löe 1963) through a non-invasive (no probing) and resetting the rating for mild and moderate inflammation (Bessa Rebelo et al. 2011). The severity of gingival inflammation is recorded from the buccal and lingual aspects of all evaluable teeth as follows:

Score 0 = absence of inflammation

Score 1 = mild inflammation or with slight changes in colour and texture but not in all portions of gingival marginal or papillary

Score 2 = mild inflammation, such as the preceding criteria, in all portions of gingival marginal or papillary

Score 3 = moderate, bright surface inflammation, erythema, oedema and/or hypertrophy of gingival marginal or papillary

Score 4 = severe inflammation: erythema, oedema and/or marginal gingival hypertrophy of the unit or spontaneous bleeding, papillary, congestion or ulceration

Gingival units as well as the calculation of the index follow the same criteria described in GI (Bessa Rebelo et al. 2011). This index was introduced for the purpose of providing an increased sensitivity in the lower degree of inflammation and to eliminate the bleeding on pressure component. Since bleeding is not a criterion, intra- and inter-examiner calibrations and reproducibility test can be performed at the same patient visit (Egelberg 1999).

18.1.10 Gingival Index (Suomi and Barbano)

The gingival index described by Suomi and Barbano (1968) as modified by Suomi (1969) is based on a combination of criteria. As an example, score 1 in Suomi and Barbano's index is based on changes in colour, volume, and texture, as well as presence or absence of stippling. The diagnostic criteria are:

Score 0 = absence of inflammation—gingiva is pale pink in colour and firm in texture. Swelling is not evident and stippling usually can be noted.

Score 1 = presence of inflammation—a distinct colour change to red or magenta is evident. There may be swelling, loss of stippling, and the gingiva may be spongy in texture.

Score 2 = presence of severe inflammation—a distinct colour change to red or magenta is evident. Swelling, loss of stippling, and a spongy consistency can be noted. There is either gingival bleeding upon gentle probing with the side of an explorer or the inflammation has spread to the attached gingiva.

18.1.11 Gingival Bleeding Index (GBI) (Ainamo and Bay 1975)

The use of the gingival bleeding index (Ainamo and Bay 1975) is simple. A blunt probe is used for gentle probing of the orifice of the gingival crevice. No pain should be caused by the probing. If

bleeding occurs within about 10 s after testing, a positive finding is recorded. The number of positive findings is then expressed as a percentage of the number of gingival margins examined. The result of clinical examination should correspond to the amount of bleeding provoked through cleaning of the teeth by the patient. It has been shown that the scores obtained using the GBI correlate statistically highly significantly to the Silness and Loe (1963) GI index scores of the same persons. The GBI has been used with success in both epidemiological profile studies and short-term clinical trials (Ainamo and Bay 1975).

In addition to the above-mentioned indices, numerous studies have used bleeding on probing (BOP) as one criterion of periodontal status without necessarily specifying the index employed. These studies have differed from the gingival bleeding index (GBI) originally described by Ainamo and Bay (1975) in that probing has usually been done at a different angulation, to a depth where resistance is encountered at the bottom of the gingival sulcus, whereas the GBI uses 'gentle probing of the orifice or the gingival crevice'. In other words, the GBI was intended to estimate gingivitis, whereas BOP probes to the bottom of a pocket and is one measure of periodontitis (Newbrun 1996).

Bleeding on probing and the gingival index have been used to clinically characterize the degree of gingival inflammation. It is, however, unclear to what extent these parameters correlate to each other and to probing pocket depth (PD). Chaves et al. (1993) evaluated the association between BOP and GI bleeding, as well as the relationship of these variables to PD, in a group of patients presenting with naturally-occurring gingivitis, using a standardized pressure sensitive probe (Florida Probe) for BOP and PD measurements. In this population, means of 40.9% (S.E. = 1.36) BOP sites and 35.3% (S.E. = 1.81) GI bleeding sites per patient were found. When sites were evaluated, BOP demonstrated a positive correlation with PD, whereas GI bleeding correlated with PD. For sites characterized by the absence of BOP as well as the absence of GI bleeding (scores 0 and 1), the highest % of agreement between the 2 indices (77.7%) was found in shallow sites (0.1–2 mm) (Chaves et al. 1993).

Quantification type indices require a much greater expenditure of time and effort in gathering the data, and the statistical analysis of data derived from such indices is vastly more complex and difficult to perform reliably than the simpler indices which rely on a binary code (i.e., 0 or 1). Thus, binary indices are not only quick and easy to administer but are also highly compatible with algorithms forming the basis of statistical packages generally in use for the analysis of clinical data (Galgut 1999).

Therefore, both the binary and the quantification indices are effective in determining clinical variables and changes in clinical variables like dental plaque and gingivitis in clinical trials. Due to the ease of application, the simplicity of the data and the reduced clinical time to obtain the data binary indices have major practical advantages. A major advantage of the binary indices when used in clinical practice is that the data is usually presented in an easily understood form. A reduction in plaque of 70–30% using a binary index is easier for patients to comprehend than a reduction of 2.21–1.45 using a quantification index. Binary/dichotomous indices can therefore be used as a powerful motivational tool in implementing an oral hygiene programme. Being able to set a target percentage or percentage reduction in oral cleanliness may provide the key element in an effective patient motivation strategy in clinical practice (Galgut 1999).

18.1.12 Gingival Bleeding Index (GBI) (Carter and Barnes 1974)

The gingival bleeding index, introduced by Carter and Barnes in 1975, records the presence or absence of gingival inflammation as determined by bleeding from interproximal sulci. Unwaxed floss was used as an index instrument, providing a quickly evaluation of a larger area of the sulcus. To score a susceptible area, the unwaxed floss was alternately passed interproximally into the gingival sulcus on both sides of the interdental papillae, and with the floss extended as far as possible towards the buccal and lingual, the floss was carried to the bottom of the sulcus. The floss was then moved in an inciso-gingival motion for one double

stroke. Care is taken not to cause laceration of the papillae, and a new length of clean floss is used for each interproximal unit. The mouth was divided into six segments and flossed in the following order: upper right, upper anterior, upper left, lower left, lower anterior, and lower right. With the examiner also retracting the cheek during the application of the floss, direct vision may be used to identify units of haemorrhage from both buccal and lingual surfaces. Bleeding is generally immediately evident in the area or on the floss; however, 30 s was allowed for reinspection of each segment. No attempt was made to quantify the degree of bleeding; only the absence or presence of bleeding from each unit is recorded. From each patient, a gingival bleeding score is obtained by noting the total units of bleeding and the total susceptible areas at risk (Carter and Barnes 1974).

Mariath et al. (2008) evaluated flossing as a diagnostic method for interproximal gingival bleeding in children. Examinations comprised repeated measurements of two gingival indices with a 10-min interval in the following sequences: the Ainamo and Bay gingival bleeding index (GBI) followed by the Carter and Barnes flossing index (CBI), CBI followed by GBI, and GBI followed by GBI. Percentage agreements in sequences GBI-CBI, CBI-GBI, and GBI-GBI were 70.3%, 76.4%, and 84.5%, respectively. Validation of flossing in the first sequence (GBI-CBI) resulted in values of 0.61 (95% CI: 0.53–0.68), 0.72 (95% CI: 0.69–0.76), 0.33 (95% CI: 0.28–0.39), and 0.89 (95% CI: 0.86–0.92), respectively, for sensitivity, specificity, positive predictive values and negative predictive values. From these results, it can be concluded that professional flossing is a useful tool in the diagnosis of interproximal gingival inflammatory status also in children, especially in conditions of gingival health.

18.1.13 Gingival Bleeding Time Index (Nowicki et al. 1981)

Nowicki et al. (1981) has developed a clinical index based on gingival bleeding time. A probe was inserted into the sulcus of the six representative areas (mesio-buccal aspect of the maxillary

right first molar, left central incisor, left first bicuspid, mandibular right first molar, right central incisor, and right first bicuspid) until light resistance was felt. The probe was then used to gently stroke the inner aspect of the gingival for one back and forth motion over a distance of approximately 2 mm. The probe was removed immediately, and both pocket depth and the time for bleeding to occur were recorded with a stop watch. If no bleeding was evident after 15 s, the stroking procedure was repeated and bleeding recorded up to an additional 15 s. If any area that previously had been assigned a score of 0 or 1, demonstrated bleeding within the 15 s after the first stimulation by gentle stroking with probe, a score of 2 was given.

18.1.14 Eastman Interdental Bleeding Index (EIBI)

Caton and Polson (1985) developed the Eastman interdental bleeding index (EIB). To test for interdental bleeding, a wooden interdental cleaner was inserted between the teeth from the facial aspect in such a way as to depress the interdental papilla 1–2 mm. The path of insertion was horizontal, taking care not to direct the point apically. In this manner, the cleaner was inserted and removed four times. The presence or absence of bleeding within 15 s was recorded. Scores are as follows: 0: absence of bleeding when a triangular toothpick is depressed 2 mm, inserted interproximally four times, and checked 15 s later and 1: bleeding after above procedure. Total number of areas that bleed are recorded (Caton et al. 1989; Lavigne 1998).

Bliesen et al. (1992) investigated the intra- and inter-examiner reliability of the method of stimulation for bleeding used in the Eastman interdental bleeding index in 26 subjects examined twice, 1 h apart, by either a single examiner or two examiners in each half of their mouths, for the presence bleeding after stimulation with a wooden interdental cleaner. Overall, intra-examiner agreement statistics were high (91.3–93.1% agreement; 0.79–0.86 kappa coefficient). Further breakdowns of the data into facial and lingual

sites by arch and location (anterior or posterior) resulted in similar levels of reliability, with no significant differences within examiners. The overall inter-examiner agreement statistics were good (82.8–87.6% agreement; 0.62–0.75 kappa coefficient). When inter-examiner data were analyzed at facial or lingual sites by arch and location, a significant difference existed in reliability for mandibular posterior lingual sites, but reliability was high in all other areas. The authors concluded that their data demonstrated a high level of reproducibility for this method, which suggests that the Eastman interdental bleeding index was suitable for clinical trials and epidemiologic studies of interdental gingivitis.

18.1.15 Periodontal Pocket Bleeding Index (PPBI) (Van der Velden 1979)

Van der Velden (1979) assessed the gingival condition by means of a periodontal pocket bleeding index (1979). The criteria were 0=no bleeding of the pocket after probing with a force of 0.75 N (1 g=1 pond=0.0098 N) and 1=bleeding of the pocket within 30 s after probing with a force of 0.75 N. The periodontal pocket bleeding index (PPBI) has been introduced in an attempt to evaluate both pocket depth and gingival health.

18.1.16 Quantitative Gingival Bleeding Index (QGBI)

In 1985, Garg and Kapoor formulated a quantitative gingival bleeding index. This index takes into consideration the magnitude of blood stains covering tooth brush bristles on brushing and squeezing gingival tissue units in a segment, with one score for entire one segment (canine to canine or left or right premolars and molars in maxillary or mandibular arches—six segments in all) (Bessa Rebelo et al. 2011). The criteria scores are:

0=no bleeding on brushing; bristles free from blood stains

1 = slight bleeding on brushing; bristle tips stained with blood

2=moderate bleeding on brushing; about half of bristle length from tip downwards stained with blood

3=severe bleeding on brushing; entire bristle length of all bristles including brush head covered with blood

Bleeding is generally immediately evident on the bristles of the brush; however, 30 s were allowed for reinspection of each segment. According with authors, this index has good reproducibility, reliability, objectivity, and simplicity of use (Bessa Rebelo et al. 2011).

18.1.17 Bleeding Index (BI)

For recording the bleeding index (modified index of Saxton), bleeding tendency was assessed after air drying by inserting the probe into the gingival crevice to a depth of approximately 1 mm and moving it around the crevice, at an angle of approximately 60° to the long axis of the tooth, stroking the inner surface of the sulcular epithelium. Bleeding detected within 30 s of probing was recorded. A score of 1 or 0 was assigned according to Löe (1967) to define a score of 1. A further score of 1 or 0 was assigned according to whether the site exhibited bleeding on probing. The data was combined to form a version of the gingival index (GI) (Löe 1967) and an inflammation index (II) (Saxton and van der Ouderaa 1989).

In a recent study performed by Newby et al. (2011), the bleeding index (modification to the index of Saxton) was performed by a single examiner using a colour-coded periodontal probe. The probe was engaged approximately 1 mm into the gingival crevice. A moderate pressure was used whilst sweeping from interproximal to interproximal along the sulcular epithelium.

The BI scoring system was as follows:

0=no bleeding after 30 s

1=bleeding upon probing after 30 s

2=immediate bleeding observed (Newby et al. 2011)

The difference between the indices was that in the inflammation index, sites which bled on probing but were without visual symptoms of inflammation were considered to be inflamed.

The gingival index created was a simplified version of the original (Löe 1967), being weighted towards bleeding tendency between scores 1 and 2. The clinical assessments were performed on the mesial, buccal, distal, and lingual aspects of the experimental teeth (Saxton and van der Ouderaa 1989).

18.1.18 Oral Rating Index (ORI) (Kawamura et al. 2000)

An oral rating index (ORI) was developed as a simple scoring system of gingival health care and oral hygiene level for adults. The ORI is based upon examination of four areas: the facial surfaces of the anterior teeth and the lingual surfaces of the right posterior teeth in the mandible and maxilla. Each patient's gingival health care level is recorded as a composite index, which categorizes gingival status and the presence and extent of local irritants (dental plaque and dental calculus) on an ordinal scale from -2 (very poor) through to +2 (excellent). When in doubt, the examiner assigns the lesser score. The assessment of the ORI score is weighted by mainly gingival condition, followed by calculus accumulation and plaque accumulation. The ORI uses a set of standard colour photographs of each level of the scale to maintain consistent standards. It is undertaken without hand instruments and using a natural light. It takes around 10 s to carry out the examination for each patient. The ORI is not a strict quantitative index for the determination of oral health status or clinical assessment. Although screening by ORI does not make a specific diagnosis of a periodontal condition, it appears helpful for identifying suspected gingival inflammation and level of oral hygiene need, without any extensive instrumentation or additional aide. It does provide a good indicator of a person's commitment to self-care. As such, it appears to be a useful public health tool to classify a person's effective oral health behaviour, rather than relying on their stated beliefs, perceptions, or practices (Kawamura et al. 2000). Scoring and criteria for the oral rating index (ORI) are as follows:

- Score (+2): good, healthy appearance; normal gingiva and no detectable plaque or calculus=excellent status
- Score (+1): clean appearance; slight localized inflammatory changes; fairly good oral hygiene=good status
- Score (0): questionable appearance; difficult to assign a positive or negative score in 10 s=questionable status
- Score (−1): poor appearance; overt gingivitis and poor oral hygiene=poor status
- Score (−2): extremely poor appearance; severe gingivitis and very poor oral hygiene=very poor status

The validity and reproducibility of the ORI as a screening method have been demonstrated by comparisons with Jackson's gingivitis index (GI), Greene and Vermillion's debris index (DI) and calculus index (CI), probing pocket depth (PPD), and bleeding on probing (BOP). The ORI was found to be significantly correlated with the GI, DI, CI, and PPD and had moderate to good intra- and inter-examiner reliability (Kawamura et al. 2000).

Using the ORI criteria, computer software (in both Japanese and English versions) named Development of Ability to Assess Gingival Status (DAAGS) was developed by one of the authors (Kawamura). He used Visual Basic for Applications (VBA) to improve dentists', dental students', and dental hygienists' ability to assess the gingival health level and/or to examine the inter- and intra-examiner reproducibility while using the ORI. VBA is a programming language that allows users to programme macros to accomplish repetitive or complex tasks automatically within Excel. The DAAGS software (Japanese version) consists of two courses: group and private. The DAAGS contains three subcourses: (1) the beginners' course for youth, (2) the advanced course for oral health instructors, and (3) the superlative course for researchers. It has been showed that the DAAGS software can serve as a useful instructional tool for education in both undergraduates and graduated dentists (Camgoz et al. 2008, 2011).

In a recent study, the ability of the second-year students in a junior high school was compared with the fourth-year dental students for

the evaluation of the gingival health status by using the DAAGS programme for beginners. The dental students group was found significantly ($P<0.05$ – 0.001) superior to the junior high school group in the ability to evaluate gingival health status.

Okada et al. (1999) developed an oral rating index for children (ORI-C) as a system for screening children's gingival health and oral hygiene status. The system has the following characteristics: (1) it requires a very short time to make a judgement (approximately 10 s), (2) it is easy to understand by children and school teachers, (3) it is readily available for screening gingival health status, and (4) it uses standardized colour photos. This scoring system is recommended for use in large-scale population surveys and as an educational and motivational method to improve children's gingival health status (Okada et al. 2000, 2002, 2008).

18.1.19 Bleeding on Interdental Brushing Index (BOIB)

Whereas measures of gingival inflammation through indices of bleeding with polling can be influenced by factors such as angulation of the probe, the probe insertion depth, direction, and motion of the probe and probing force and indices that use wooden spatulas, according to its shape and rigidity, may represent a potential for trauma, Hofer et al. (2011) developed the bleeding on interdental brushing index (BOIB) (Bessa Rebelo et al. 2011). This index is scored by passing through with a light interdental brush (Curaprox CPS Prime; Curaden AG, Kriens, Switzerland) (Fig. 18.3) placed buccally, just under the contact point and guided between the teeth with a jiggling motion, taking care not to exert force. If the brush met any resistance, a smaller brush was substituted and the insertion procedure was repeated. Bleeding was scored as either present or absent, for each interdental site within each quadrant, after 30 s (Fig. 18.4). The advantages of using an interdental brush to test for bleeding include atraumatic manipulation of the papillae, ease of application, integration into existing oral hygiene

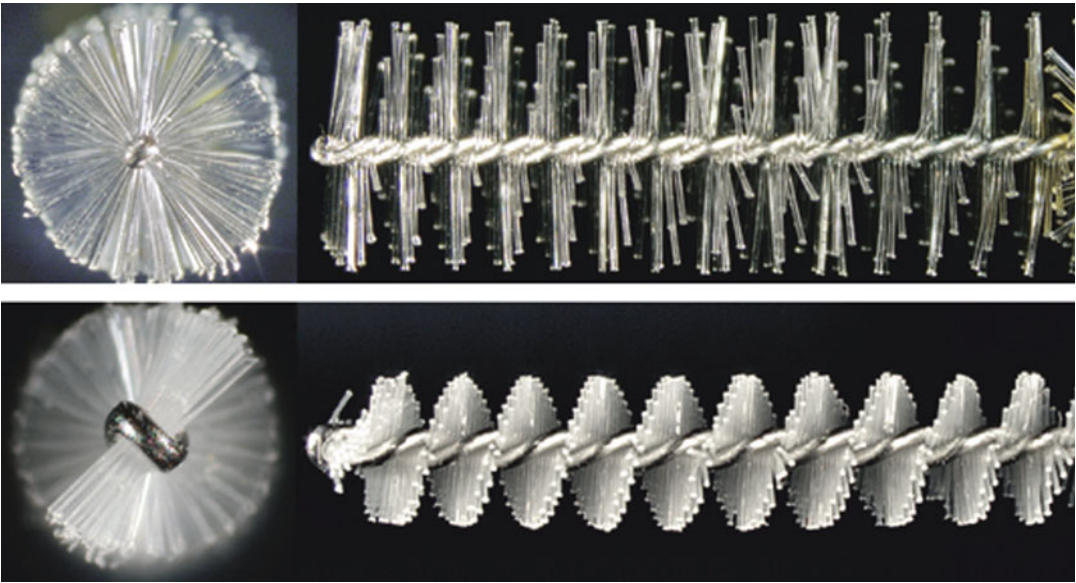


Fig. 18.3 Upper row: a ‘light’ interdental brush (CPS prime 109, wire diameter 0.9 mm, outer brush diameter 4.0 mm) used for the bleeding on interdental brushing (BOIB) assessment. Lower row: a standard interdental brush (CPS regular 110, wire diameter 1.0 mm, outer brush diameter 2.2 mm). Note the denser, more rigid filaments (Hofer et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

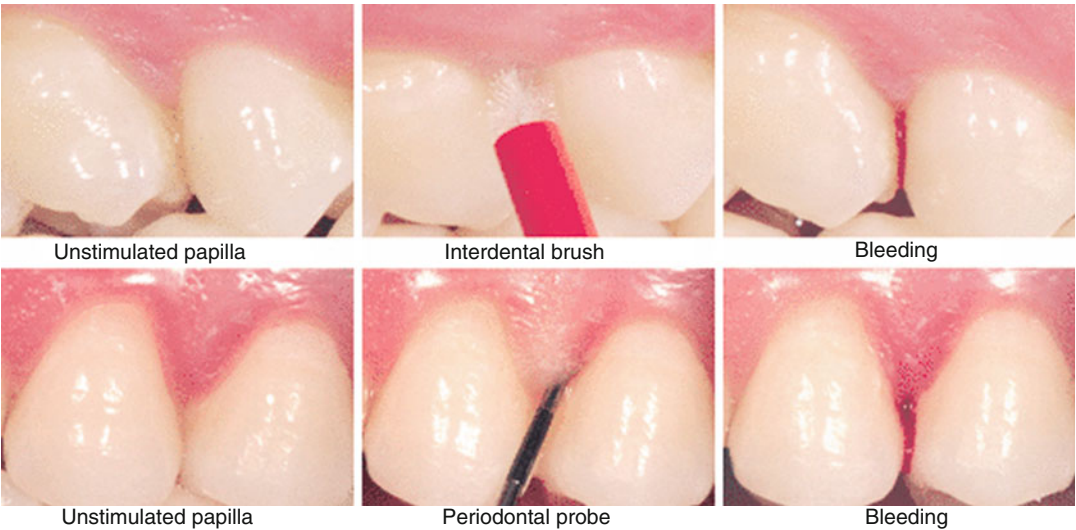


Fig. 18.4 Upper row: bleeding on interdental brushing (BOIB). Lower row: bleeding on marginal probing (BOMP) (Hofer et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

instruction, and motivating patients to monitor their own progress at home, while at the same time performing a beneficial oral hygiene procedure and removing any interdental plaque that may be present (Hofer et al. 2011).

18.1.20 Bleeding on Marginal Probing (BOMP)

Van der Weijden and co-workers (1994a) have assessed gingival bleeding on marginal probing

by inserting the probe into the gingival crevice to a depth of approximately 2 mm and furthermore running the probe along the marginal gingiva, at an angle of approximal 60° to the longitudinal axis of the tooth. Presence or absence of bleeding was scored within 30 s after probing. This method was compared to probing with a periodontal probe to the bottom of the pocket. Bleeding on marginal probing was found to be the most appropriate method to detect differences in the development of gingivitis between experimental groups (Van der Weijden et al. 1994a, b). The ability of the BOMP index to assess the level gingival inflammation appears to be comparable with the Eastman interdental bleeding (EIB) index (Barendregt et al. 2002).

18.1.21 Maintenance of Gingival Health Index (MGHI) (Butler et al. 2011)

Butler et al. (2011) investigated a method to measure maintenance of gingival health and is based on the score at each tooth site assessed and if it improves, shows no change, or worsens in score compared to a reference visit, for example, baseline. A maintenance of gingival health index (MGHI) is constructed for each subject based on all sites assessed in the whole mouth based on these three outcomes. The MGHI can be calculated for an individual treatment and compared between treatments using conventional statistical techniques.

The MGHI was calculated for both MGI (modified gingival index) (Lobene et al. 1986) and BI (bleeding index) (Saxton and van der Ouderaa 1989) separately and was derived from tooth site level changes compared to a reference visit.

MGHI was defined as $MGHI = (NDEC + NNOCHNG) / NINC$,

Where:

- NDEC = the number of sites that show a decrease in their index score (visit score—reference visit score < 0). This is seen as an improvement in score.
- NNOCHNG = the number of sites that do not change in their index score (visit score—reference visit score = 0).

- NINC = the number of sites that increase in their index score (visit score—reference visit score > 0). This is seen as a worsening in score.

This index is analogous to an odds of the maintenance of gingival health such that if the MGHI is > 1, this means that there is a greater number of sites that do not increase compared to the number of sites that do increase, that is, a greater maintenance of health. The ratio of MGHI for treatment A/MGHI for treatment B is defined as the MGHI ratio which in essence is an odds ratio. If the MGHI ratio is > 1, this translates to the odds of maintaining gingival health being greater on treatment A compared to treatment B (Butler et al. 2011).

If the number of sites that worsen is zero, the calculation of the MGHI would not be possible due to division by zero. To avoid this, the MGHI can be set to the total number of sites assessed in the subjects' mouth, and this would still maintain the ordered nature of the measure. If a subject had 108 evaluable sites and only one worsened, then the MGHI would be $107/1 = 107$. If zero sites had worsened, then we would set the MGHI to the number of sites assessed, which would be 108. This is numerically larger than its nearest neighbour but more meaningful than if infinity ($108/0$) is used. MGHI can be calculated separately for MGI and BI scores. So for each subject, an MGHI value is constructed which ranges from 0 to 108. This lends itself to be analyzed using an analysis of covariance (ANCOVA). The statistical model used for the two example studies included factors for treatment group and the pre-prophylaxis baseline and randomization baseline levels of MGI and BI (whole mouth) as covariates (Butler et al. 2011).

The MGHI is better than just looking at, for example, the proportion maintained as it encompasses all aspects of the change in index score at the site in that the score is maintained or better or it worsens in score. It is also for this same reason that the MGHI is more advantageous than just looking at the mean score, as a single site change can have a greater impact on the score than on an overall mean. This can be seen from Fig. 18.5 as the MGHI changes quite dramatically the closer

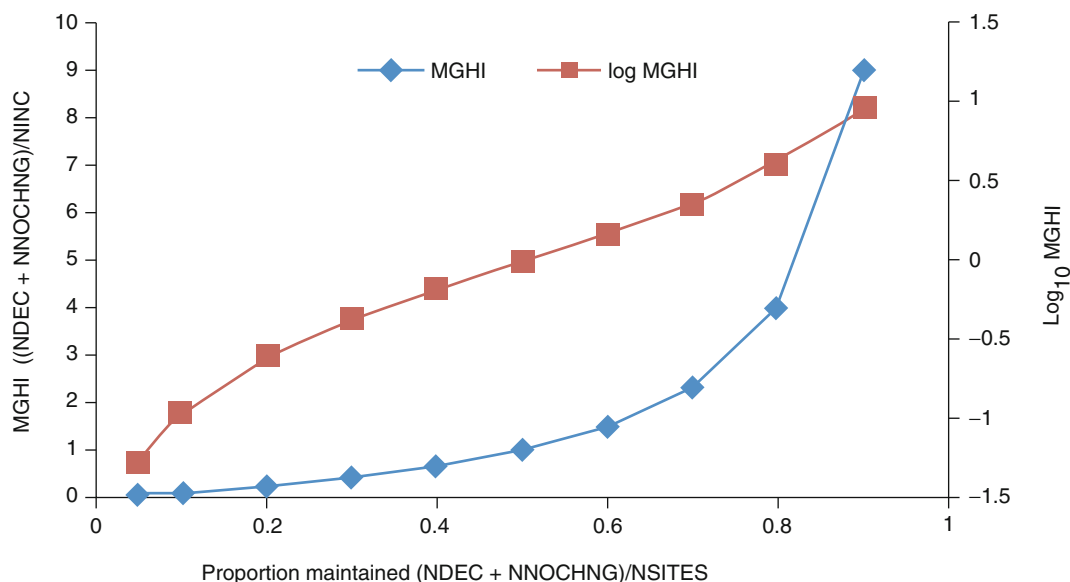


Fig. 18.5 *NDEC* the number of sites that show a decrease in index score, *NNOCHNG* the number of sites that do not change in index score, *NINC* the number of sites that increase in index score, *NSITES* the number of tooth sites.

Plot to show MGHI against the proportion maintained (Butler et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

to 100% of the sites are maintained. The whole mouth mean would be constrained between a score of 0–4 for the MGI. This single site impact can be important from a subject's viewpoint as they could suffer an oral health problem at a single site which might require a visit to a dental practice. This increased sensitivity at this end of the scale is an advantage of the MGHI over an overall whole-mouth mean score. It is also better than looking at the overall mean as increases and decreases can cancel each other out. If 50% of the scores increase by 1 point and 50% of the scores decrease by 1 point, this would lead to an overall mean which is the same. Looking at no change in the whole mouth means would indicate that close to 100% of sites are maintained. In contrast, using the individual site changes would be closer to 50% (Butler et al. 2011).

18.1.22 Imaging Techniques to Assess Gingivitis

Recently, researchers have evaluated the use of imaging techniques to assess gingivitis. Methods,

such as the photographic system produced by Ellis et al. (2001), used photographs of patient's teeth that were projected onto a large screen. These were assigned a score indicating the encroachment level of the anterior papilla from 0 to 3. This method was applied to the labial surfaces of the anterior teeth. The scoring system was based on that developed by Seymour et al. (1985) whereby the score 0 indicated no health problems, 1 defined slight granulation of appearance of the papilla, 2 indicated slight gingival encroachment over the tooth surface less than a quarter of the tooth width, and 3 described severe encroachment over a quarter of the tooth width. This method used basic image analysis but relied upon the three-stage scoring applied to each of 10 papillae regions (five upper and five lower). The limitations of most indices are related to the subjectivity of examiner scoring and the level of method standardization achievable (Smith et al. 2008).

Another method for assessing gingival changes, described by Smith et al. (2008), assessed redness and tooth surface area visible between the level of the interproximal papillae and the gingival margin as follows. After image acquisition using Adobe

Photoshop (V5.02; Adobe Systems Ltd, San Jose, CA, USA), images were measured as shown in **Figs. 18.7** and **18.8**. A thresholding process automatically selected a predetermined range of pixel values of red from the total 256 available. This

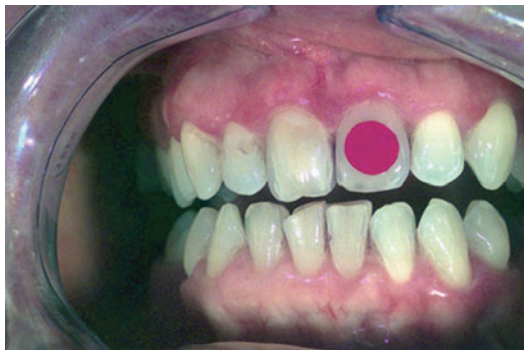


Fig. 18.6 Image of the anterior teeth of a patient with gingival inflammation showing a red calibration disc attached to the upper left central incisor (Smith et al. 2008. Reprinted with permission from John Wiley & Sons, Inc.)

process was used to isolate a red disc of known size (**Fig. 18.6**, incorporated in the images for colour and size calibration) and then the total gingivae visible on the image: The histogram option was used to give the mean red pixel value of each (**Fig. 18.7**) which was then recorded for later subtraction from the red disc value. Calibration and analysis of images was carried out using Image Pro Plus software (V4.0; Media Cybernetics, Silver Spring, MD, USA). This process was repeated on the post-treatment images gained at the patients' subsequent visits. Any slight differences in the image illumination or in patient position would be accounted for by always subtracting the standard red-coloured disc mean red pixel value from that of the gingivae, as the mean value of the red disc should not change in ideal conditions. This method ensures that changes seen are due to gingival changes not random or systematic errors (Smith et al. 2008). The major limitation of

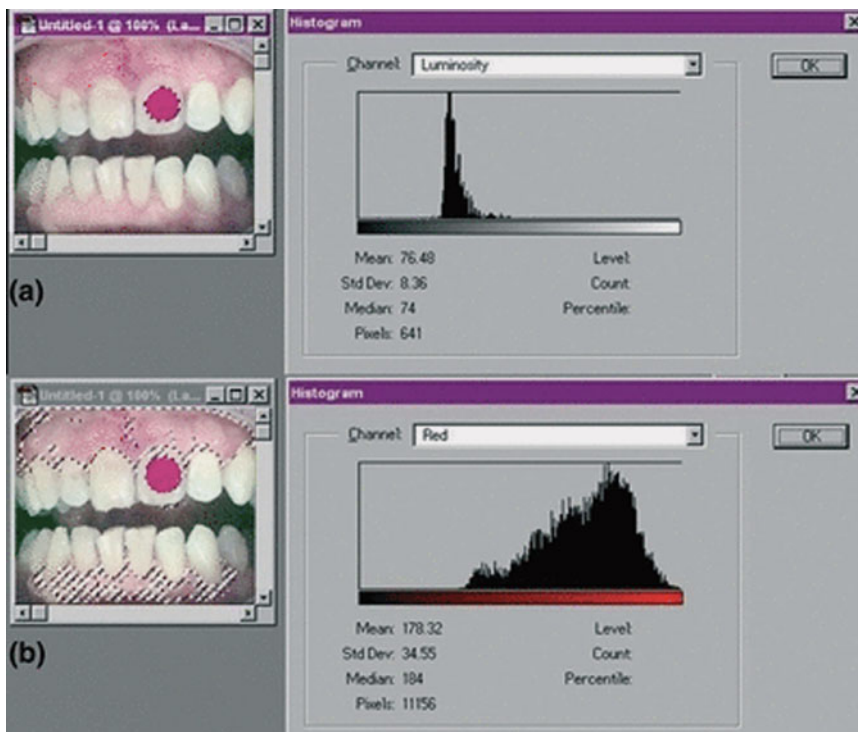


Fig. 18.7 (a) Image of the isolated red calibration disc and mean red pixel output menu; (b) image of the visible gingival area isolation by thresholding and area measure-

ment is shown in the output menu after calibration (Smith et al. 2008. Reprinted with permission from John Wiley & Sons, Inc.)

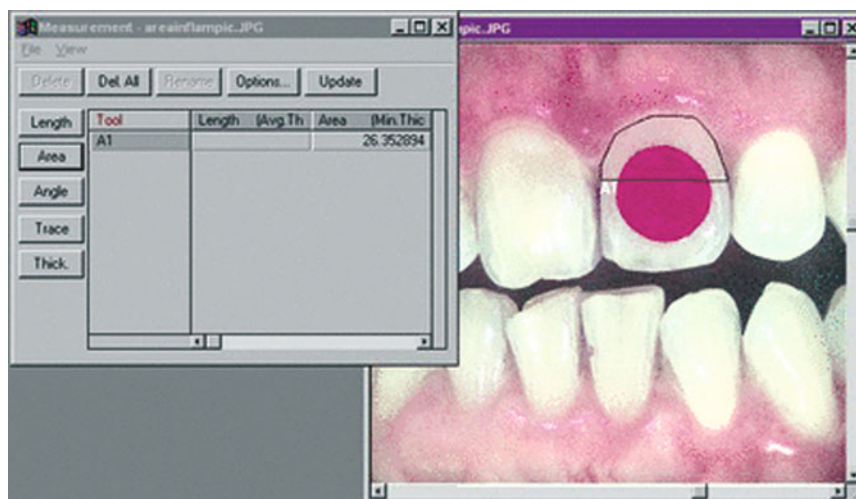


Fig. 18.8 Image of teeth showing portional tooth surface area measurement (mm^2) of an *upper* anterior tooth and the output menu after linear calibration (Smith et al. 2008. Reprinted with permission from John Wiley & Sons, Inc.)

this technique is that it only includes the assessment of the facial surfaces of the anterior teeth. When compared to the entire mouth, this accounts for only a small proportion of the sites which are treated and does not allow for assessment of those areas of the mouth that have higher levels of disease and are more difficult for the patient to clean to achieve changes in the gingival condition (Newby et al. 2011).

Rosin et al. (2002) investigated the suitability of measuring volume differences in gingival papillae for monitoring changes in the inflammatory status of the gingivae. To measure gingival oedema, impressions were taken, and the buccal side of the replicas, comprising the buccal and adjacent occlusal tooth surfaces, the papillae, and the marginal gingiva, was scanned with a 3D laser scanner (Laserscan-3D pro, Willytec, München, Germany). To calculate the papillary volume difference, the papillary region was delimited as a polygon on the image taken in the swollen condition by drawing a connecting line between the two mid-buccal points of the two adjacent teeth and marking the sulcus line to complete the polygon

(Fig. 18.9). After calculating the difference image of the two superimposed scans, the polygon is imported and cut out (Figs. 18.10 and 18.11). The volume and base of the papilla defined by the polygon can be read off and exported from the statistics module of the match programme (Fig. 18.11). A potential source of error in the Rosin study is introduced during the preparation of the impression and replicas. The volume changes observed are relatively small; therefore, poor quality impressions could have a significant impact on the data (Newby et al. 2011).

A new method for quantifying the gingival contour volume primarily designed for in-office use used the Lava™ Chairside Oral Scanner (COS) (3M), a digital impression system that replaces traditional impression materials and utilizes continuous 3D video images to create a digital impression. The teeth and gingivae were first scanned to create a digital impression, and the 3D coordinates from the digital impression were then transferred to specialized software to directly compare changes in gingival volume. Data from the hard tissue (teeth) was used to

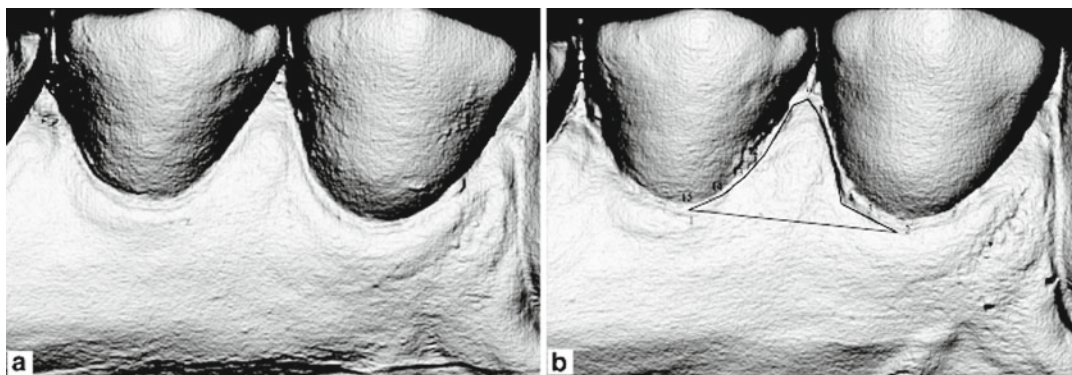
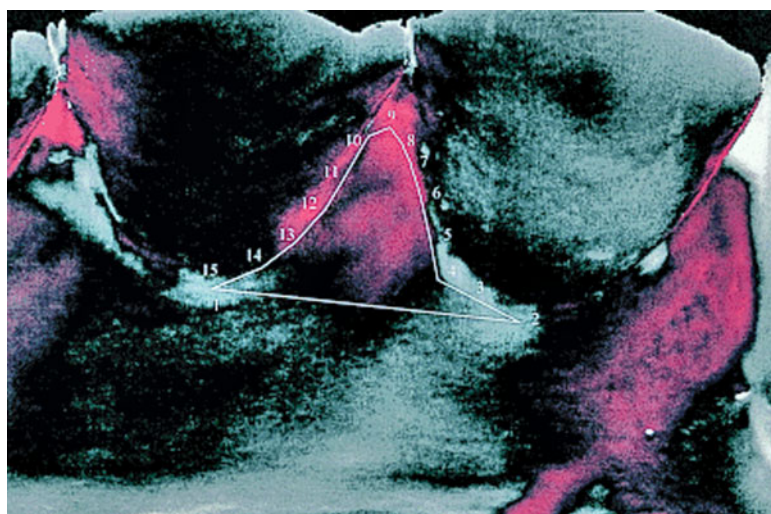


Fig. 18.9 Scanned image of a papilla between FDI teeth 4 and 5 on day 0 (**a**) and on day 28 with polygon delineating the papilla (**b**). Note that the detail and scanning angle of (**a** and **b**) are not identical. The software's algorithm uses

identical image regions (e.g., plaque-free tooth surfaces) to superimpose the two 3-D data sets (Rosin et al. 2002. Reprinted with permission from John Wiley & Sons, Inc.)

Fig. 18.10 Difference image from **Fig. 18.9a, b**. The detail corresponds to **Fig. 18.11**. The *polygon* marked in **Fig. 18.9b** was transferred into the difference image. The *red colour* indicates a positive value in the z-axis of the difference image, the brighter the colour, the higher the value (Rosin et al. 2002. Reprinted with permission from John Wiley & Sons, Inc.)



align the images from the same subject at baseline and subsequent visits as any changes that occurred over this study time would be minimal. Indeed, from the comparison data, this was the case with most teeth showing little or negligible deviations from baseline. Once the hard tissue alignment was finalized, the difference between the superimposed baseline and each subsequent

visit image was measured as the average deviation (d) in millimetres over a specific area around the gingival margin and papilla for each tooth (facial and lingual aspects). Average deviations were calculated over circular areas known as bomb radii (BR) so that error due to rogue points of scan data was minimized. Multiple BR were used to cover the entire region as illustrated in

Fig. 18.11 The papilla delineated by the polygon is cut out of the difference image. The volume (total volume) and base (n -defined) of the papilla can be read from the statistics module of the match programme and exported (Rosin et al. 2002. Reprinted with permission from John Wiley & Sons, Inc.)

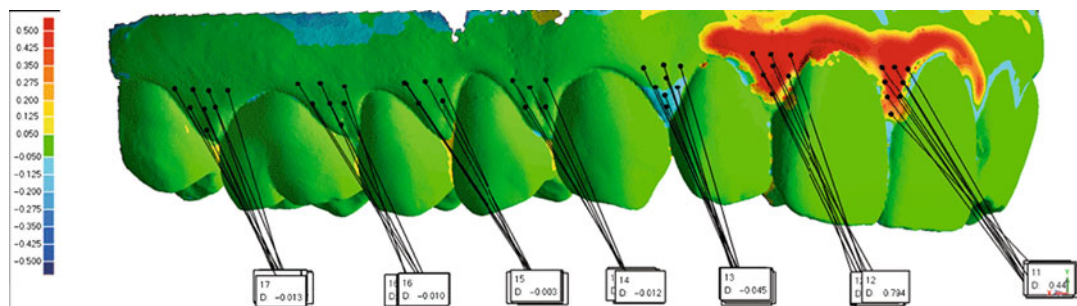
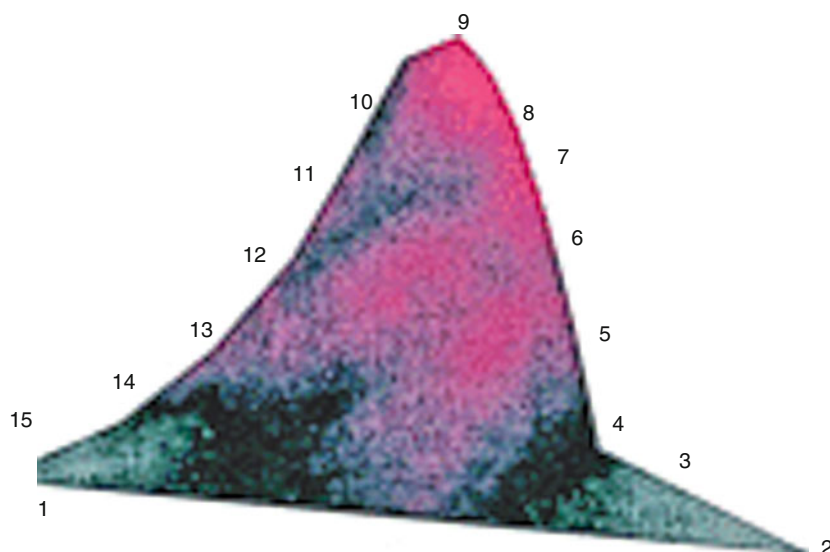


Fig. 18.12 Example of bomb radii used to calculate gingival volume changes—each single black dot represents a bomb radii (Newby et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

Fig. 18.12. Images of example changes in gingival contour scans for one subject (subject 10) are provided in **Fig. 18.13** (Newby et al. 2011).

The relationships between gingival contours and traditional indices variables are depicted in **Fig. 18.14**. There was some correlation in change in gingival contour volume against the change in modified gingival index MGI (Lobene et al. 1986) ($\rho=0.36$) and bleeding index (Saxton et al. 1989) ($\rho=0.33$). In general, there is a corresponding increase in gingival contour volume with an increase in MGI and BI (Newby et al. 2011).

The benefits of this technique are that it is quantitative and does not rely on examiner interpretation of a nonlinear index. This technique enables assessment of the entire mouth unlike previously reported techniques in digital photography which are significantly limited by access to only facial surfaces of the anterior teeth. Potential drawbacks of this technique are that it requires specialized equipment and software and the scanning process and data analysis are time-consuming. In order to assess the subjects' condition, other parameters such as colour, texture, and bleeding should be included (Newby et al. 2011).

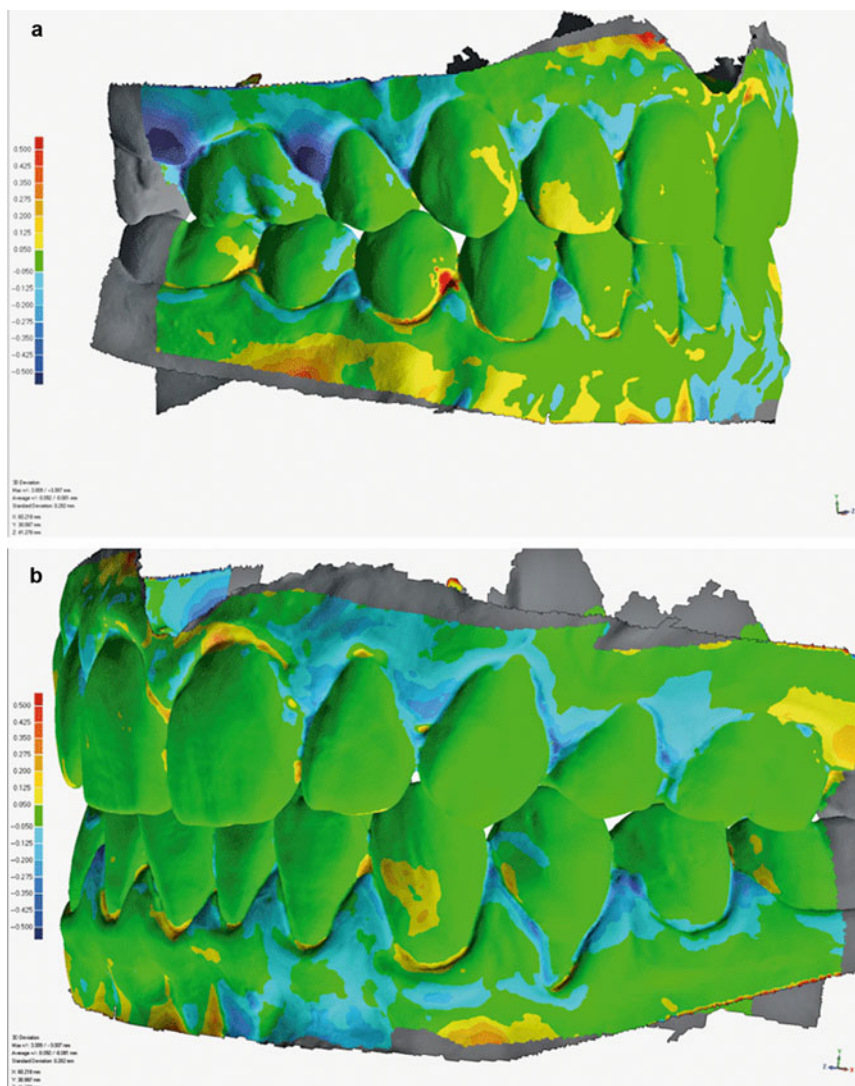


Fig. 18.13 Change in contour volume from baseline week 1 (a) *right side, upper prophylaxis*, (b) *left side, lower prophylaxis* (Newby et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

18.2 Indices for Measuring Periodontitis

The formulation of the various indices has been rooted in the prevailing understanding of the aetiology and progression of periodontal

disease at the time. The following passages provide an appraisal along with the shortcomings of indices used for measuring periodontitis, in chronological order, which is also summarized in **Table 18.1** (Dhingra and Vandana 2011).

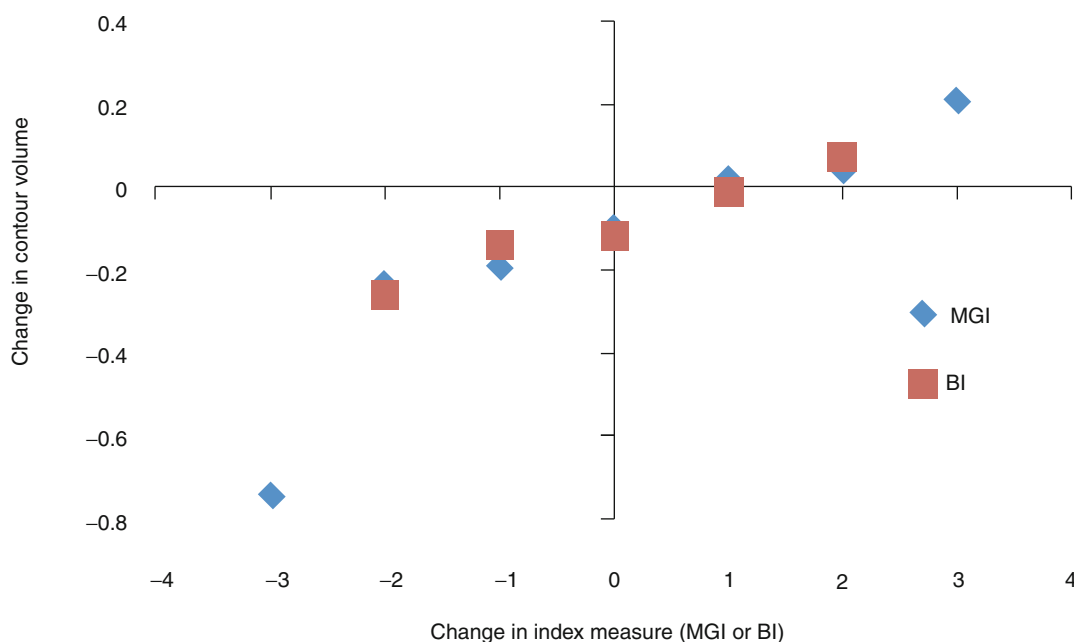


Fig. 18.14 Plot of change in MGI or BI relative to change in contour volume (Newby et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

18.2.1 Russell's Periodontal Index (PI) (Russell 1956)

The periodontal conditions of individuals can be scored by means of the periodontal index (PI) system (Russell 1956). The periodontal condition is estimated individually for each tooth in the mouth and is scored according to a progressive scale, which gives relatively little weight to soft-tissue inflammation and relatively great weight to destruction of bone. The score for an individual patient is the arithmetic mean of the scores for the teeth in his mouth. The population score is the arithmetical mean of the individual scores for the persons examined. Scores may be computed for a population of teeth as well as for persons.

Criteria for field scoring are outlined below, summarizing also the X-ray criteria used in the clinical test of the method: score 0=*negative*. There is neither overt inflammation in the investing tissues nor loss of function due to destruction of supporting tissues: score 1=*mild gingivitis*. There is an overt area of inflammation in the free gingivae, but this area does not circumscribe the

tooth: score 2=*gingivitis*. Inflammation completely circumscribes the tooth, but there is not apparent break in the epithelial attachment: score 6=*gingivitis with pocket formation*. The epithelial attachment had been broken, and there is a pocket (not merely deepened gingival crevice due to swelling in the free gingivae). There is no interference with normal masticatory function; the tooth is firm in his socket and has no drifted: score 8=*advanced destruction with loss of masticatory function*. The tooth may be loose, may have drifted, may sound dull on percussion with a metallic instrument, and may be depressible in its socket.

According to Lavigne (1998), calculations are made as follows:

- Assign a score of 0–8 for mesial, distal, buccal, and lingual surfaces of all the teeth in the mouth.
- Interpret final score as follows:
 0.0–0.2 = clinically normal
 0.3–0.9 = simple gingivitis
 1.5–5.0 = established destructive periodontal disease
 3.8–8.0 = terminal periodontal disease

Table 18.1 Description of periodontal indices (Dhingra and Vandana 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Serial no.	Index	Author and year	Salient features
1.	Periodontal Index	Russell (1956)	- First index for periodontal disease • A weighted categorical scoring system - No longer considered valid
2.	Periodontal Disease Rate Index	Sandler and Stahl (1959)	Each tooth is assessed utilizing radiographs and clinical measurements
3.	Periodontal Disease Index	Ramfjord (1959)	More sensitive version of the PI for use in clinical trials • Scores the gingival status first using a 0–3 scale while clinical attachment level is scored on a scale from 4 to 6, on a selected group of teeth—'Ramfjord teeth'
4.	Gingival Bone Count Index	Dunning and Leach (1960)	Subjective measurement of gingival status is combined with proportionate measurements of bone loss from radiographs • Time consuming, thus, not used
5.	Gingival Periodontal Index	O'Leary (1967)	Mouth is divided into six sextants and the highest score for each segment either gingival (0–3) or periodontal (4–6) is recorded • Not much used
6.	Navy Periodontal Disease Index	Grossman and Fedi (1974)	Derived from gingival (0–2) and pocket (0, 5 and 8) scores of the six Ramfjord teeth
7.	Community Periodontal Index for Treatment Needs	World Health Organization by Ainamo, Barmes, Beagrie, Cutress, Martin and Infirri (1982)	Assesses the presence or absence of gingival bleeding on probing, supra or subgingival calculus and periodontal pockets by using 0.5 mm ball tip WHO probe • Advantages include simplicity, speed and international uniformity, hence, popular
8.	Periodontitis Severity Index	Adams and Nystrom (1986)	Assesses the presence or absence of periodontitis as product of clinical inflammation and interproximal bone loss determined radiographically using a Schei ruler • Use limited to longitudinal studies and lacks validation
9.	Extent and Severity Index	Carlos et al. (1986)	Extent score is % of sites examined having attachment loss more than 1 mm whereas, the severity score is average loss of attachment per site among disease sites • Simple, reproducible
10.	National Institute of Dental and Craniofacial Research (NIDCR) protocol	NIDCR (1988–1994)	Periodontal examination consisted of measurement of periodontal supporting tissues including attachment loss, probing pocket depth and furcation involvement • Used in U.S. National Health and Nutrition Examination Survey (NHANES III)
11.	Periodontal Index for Treatment	Eaton and Woodman (1989)	Clinical assessment of six teeth is done with a specially designed periodontal probe • Simple rapid and reliable periodontal screening

(continued)

Table. 18.1 (continued)

Serial no.	Index	Author and year	Salient features
12.	Periodontal Screening and Recording Index	American Academy of Periodontology (1991)	- Divides mouth into six sextants and greatest score in each sextant of mouth is determined and recorded by using a plastic PSR probe - Simple, fast and preferred by patients
13.	Community Periodontal Index	World Health Organization (1997)	- Modification of CPTN index - Useful in periodontal research, especially to reduce the time needed for examinations when the study population comprises a large number of individuals
14.	Periodontitis Index	Albandar et al. (1999)	Classifies each person as having either mild, moderate or advanced periodontitis, or with no periodontitis, based on the number (or percentages) of teeth showing certain thresholds of probing depth and attachment loss
15.	Dichotomous Periodontal Index	Dye et al. (2002a, b), Tezal et al. (2004), Borrell et al. (2006), Brothwell and Ghiabi (2009), Persson et al. (2005)	Record the presence or absence of pocket or clinical attachment loss against a cut-off point
16.	Genetic Susceptibility Index for periodontal disease	Moustakis et al. (2007)	Used for both single nucleotide polymorphisms and microbial component of periodontal disease

In 1967, Russell provided a refinement in the form of treatment needs based on the PI, that is, PI scores in the range of 0.1–1.0 require simple prophylaxis, 0.5–1.9 require minimal periodontal treatment, 1.5–5.0 require elaborate and protracted treatment, while 4.0–8.0 require full-mouth extraction (Ekanayaka and Sheiham 1979).

From a fundamental point of view, Russell's periodontal index records three crucial stages in periodontal stages in periodontal destruction: gingivitis (scores 1 and 2), pocket formation (score 6), and the almost breakdown of the periodontium (score 8). This index does not differentiate between shallow or fairly deep pockets, except at the stage when the tooth is about to lose its function. In essence, the periodontal index is a morbidity index, which merely answers yes or no to whether the tooth has gingivitis, has pocket formation, or has lost its function due to periodontal destruction. This is the strength of the periodontal index in assessing the overall periodontal disease prevalence in large population samples and its weakness when smaller samples or when the effects of preventive and therapeutic measures are to be analyzed (Løe 1967).

For a strictly epidemiological study of variations in distribution of periodontal disease in populations, the Russell index seems to be satisfactory, but the method has also some serious shortcomings (Ramfjord 1959):

1. Although Russell's periodontal index has two scores for gingivitis (scores 1 and 2), this index does not really consider different qualities of gingival inflammation. The scores for gingivitis do not refer to various degree of severity of the pathological condition but merely to the horizontal extension of the marginal inflammation around the tooth (Løe 1967). Inadequate examination procedures based on cursory inspection of the gingival tissues without drying off the saliva for evaluating colour and without palpation for density and without routine probing for pockets is a background for underestimation of the periodontal disease prevalence (Ramfjord 1959).
2. The recording of periodontal pockets by this method is glossy inadequate because of unsuitable tools, lack of routine probing for pockets, and lack of orientation of pockets to the cemento-enamel junction. With superficial inspection, a number of pockets will remain unrecognized unless radiographs are available. This error will be most significant in mouths exhibiting good oral hygiene and little overt inflammation (Ramfjord 1959).
3. Increase in mobility is not recognized until the tooth is so loose that it cannot be used for function. If mobility should be included as a factor to be recorded, it seems desirable to have some provision for recording it prior to the terminal stage of periodontal disease (Ramfjord 1959).

Presently, Russell's periodontal index is no longer considered valid as it was flawed, conceptually and methodologically, in that gingivitis is no longer considered to be the equivalent of early periodontitis and the index did not measure features specific for periodontitis (in contrast to gingivitis), such as pocket depth, clinical attachment level, and radiographic bone loss (Dhingra and Vandana 2011).

18.2.2 Periodontal Disease Index (PDI) (Ramfjord 1959)

Like the periodontal index (PI) (Russell 1956), the periodontal disease index (PDI) (Ramfjord 1959) has two components, for gingivitis and periodontitis. The areas surrounding a selected group of teeth called '**Ramfjord teeth**' (teeth number 16, 21, 24, 36, 41, and 44) are examined. The gingival areas around each of the above-mentioned teeth are observed for deviations from health in colour, form, density, and bleeding tendency. It was recommended to record gingival findings as follows:

Score G 0 = absence of inflammation

Score G 1 = mild to moderate inflammatory gingival changes not extending all around the tooth

Score G 2 = mild to moderate inflammatory gingival changes extending all around the tooth

Score G 3 = severe gingivitis characterized by marked redness, tendency to bleed, and ulceration

The individual index for the periodontal disease is obtained in the following manner: First, a periodontal score for disease is tabulated for each of the examined teeth. If the gingival crevice in none of the measured areas extends apically to the cemento-enamel junction, the recorded score for gingivitis is the score for periodontal disease for this tooth. If the gingival crevice in any of the four measured areas extends apically to the cemento-enamel junction but not more than 3 mm (including 3 mm) in any area, the tooth is assigned a score for periodontal disease of 4; the score for gingivitis for that tooth is then disregarded in the final index for periodontal disease. If the gingival crevice in any of the four recorded areas of the tooth extends apically from 3 to 6 mm (including 6 mm) in relation to the cemento-enamel junction, then the tooth is assigned a score for periodontal disease of 5 (the gingivitis score also is disregarded here). Whenever the gingival crevice extends more than 6 mm apically to the cemento-enamel junction in any of the measured areas of the tooth, the score of 6 is given as score for periodontal disease for that tooth (again disregarding the gingivitis score) (Ramfjord 1959).

The calculation is as follows (Lavigne 1998): PDI score = total scores for all teeth/number of teeth examined; significance: group score of 3.5 = severe gingivitis; and scores of 4–6 depict periodontitis.

Ramfjord's periodontal disease index is also a composite system, which records both the gingival and periodontal situations. In this system, the scores for periodontal destruction is based on loss of attachment as measured in millimetre from the cemento-enamel junction to the bottom of the pocket, allowing a more precise quantification of periodontal destruction, compared with Russell's periodontal index (Löe 1967).

Its main advantages were easy calibration of examiners, and it was an accurate tool to assess periodontal status as it was based on measurements rather than estimation. However, it was time-consuming due to a high requirement of precision, and also in primitive populations with many middle-aged and older people, much time

has to be spent on removing calculus to determine the location of cemento-enamel junction. Although the PDI was never used for national estimates of periodontal disease in the USA, it was used for various epidemiological surveys in India and Michigan and various clinical trials of therapeutic or preventive procedures (Dhingra and Vandana 2011). The revolutionary concepts of clinical attachment level and Ramfjord teeth, that is, partial-mouth recording are now being used for USA national oral health surveys (Dhingra and Vandana 2011; Dye et al. 2007; Ramfjord 1961; Chawla 1963; Mandel et al. 1997).

18.2.3 Gingival Bone Count Index (GB) (Dunning and Leach 1960)

In 1960, Dunning and Leach formulated the gingival bone (GB) count index, in which the subjective measurement of gingival status is combined with proportionate measurements of bone loss from radiographs to produce a composite score (Table 18.2). The GB count may be used as a composite index, morbidity index, or cumulative index. Since it combines clinical examination with reading of radiographs, it is one of the most time-consuming of all the methods and is thus not extensively used (Dunning and Leach 1960; Dhingra and Vandana 2011).

18.2.4 Periodontal Disease Rate (PDR) (Sandler and Stahl 1959)

Periodontal disease rate (PDR) was proposed for epidemiological studies of periodontal disease because the measurement is meaningful and clinical assessment is easy and rapid. It was stated in the form $PDR = \frac{a}{a+b}$ in which PDR is the peri-

odontal disease rate, a is the number of teeth affected by periodontal disease, b is the number of teeth not affected by periodontal disease, and $a+b$ is the tooth population of the individual patient. The authors stated that the periodontal disease rate produces similar results with the Russell's periodontal index (Sandler and Stahl 1959).

Table 18.2 Scoring system for GB count (Dunning and Leach 1960) (Reprinted with permission from Sage Publications)

Gingival score (one score is assigned to each tooth studied. A mean is then computed for the whole mouth).	
Negative	0
Mild gingivitis involving the free gingiva (margin, papilla, or both)	1
Moderate gingivitis involving both free and attached gingiva	2
Severe gingivitis with hypertrophy and easy haemorrhage	3
Bone score (one score is assigned to each tooth studied. A mean is then computed for the whole mouth).	
No bone loss	0
Incipient bone loss or notching of alveolar crest	1
Bone loss approximating one-fourth of root length or pocket formation one side not over one-half root length	2
Bone loss approximating one-half of root length or pocket formation one side not over three-fourths root length; mobility slight ^a	3
Bone loss approximating three-fourths of root length or pocket formation one side to apex; mobility moderate ^a	4
Bone loss complete; mobility marked ^a	5
Maximum possible GB count per perso	8

^aIf mobility or impairment of masticatory function varies consider

18.2.5 Periodontitis Severity Index (PSI) (Adams and Nystrom 1986)

A periodontitis severity index (PSI) was developed and then tested by Adams and Nystrom (1986). The essential premises upon which the PSI is based are (1) periodontitis is diagnosed on the concurrence of clinically apparent marginal inflammation and vertical loss of supporting periodontium and (2) a Schei ruler is used to determine from the radiograph the percentage of bone loss from a tooth surface. Periodontitis severity can be determined by calculating the percent of vertical supporting alveolar loss at an inflamed site.

For each patient, a 0–1 clinical inflammation score (CIS) was determined at the mesial and distal of every tooth. A CIS of 0 reflected no demonstrable clinical inflammation, whereas a CIS of 1 meant that clinical inflammation could be detected: oedema, suppuration, bleeding upon provocation, increased crevicular fluid flow, or colour deviation. A modified Schei ruler was used to measure the radiographs. The ruler was changed from that it originally described in that it acknowledged the typical 1.5-mm distance between the alveolar bone crest and the

cemento-enamel junction in ideal health and was divided into tenths, permitting analysis of the percentage of bone loss in 10% increments. At each mesial and distal tooth surface, a whole bone loss score (BLS) was determined using the modified Schei ruler. A BLS of 0 meant no bone loss, 1 = up to 10%, etc., up to BLS10 = 90–100% loss. Then, for each surface, a PSI was calculated using the formula:

Periodontitis Severity Index = Clinical Inflammation Score X Bone Loss Score; $PSI = CIS \times BLS$

The PSI was calculated at each mesial and distal surface, and a mean for the dentition was calculated (Adams and Nystrom 1986).

The authors stated that the PSI offers several advantages. Healthy sites can be distinguished from diseased sites, ratio data can be produced, arbitrarily weighted clinical observations can be largely avoided, and direct measurement of periodontitis can be made. The disadvantages are that in order to recalculate the PSI over time, further radiographs are necessary, and also, radiographs do not permit buccal or lingual PSI calculations (Adams and Nystrom 1986). Due to its limitations, the PSI is limited to longitudinal studies and lacks validation (Dhingra and Vandana 2011).

18.2.6 Extent and Severity Index (ESI) (Carlos et al. 1986)

The extent and severity index (ESI) (Carlos et al. 1986) has enjoyed substantial popularity because it not only establishes the severity of periodontitis as the mean attachment loss in diseased sites but also incorporates the proportions of all sites that are affected by periodontitis (extent). Under this index, in two designated quadrants, loss of attachment is determined in two periodontal sites, that is, mid-buccal and mesio-buccal of each tooth except the third molars. This results in 28 measurements for each subject. For an individual, the extent score is the percentage of sites examined that have attachment loss more than 1 mm, whereas the severity score is the average loss of attachment per site among the disease sites. The ESI for a population is the average extent and severity scores for the individuals examined. Its advantages are that it is simple, reproducible method which appears to yield informative description of periodontal disease status of a population and requires only minimal training of examiners (Carlos et al. 1986; Dhingra and Vandana 2011). One of the uses of the ESI is the comparison of extent and severity scores among subgroups of people in surveyed populations (Hunt and Fann 1991).

Because the ESI measures attachment loss in only 28 sites in contralateral quadrants, it has the same potential for inaccuracy as other partial measurements. When the extent and severity index (ESI) was compared with a gold standard represented by actual readings of loss of attachment on six sites around all teeth present (excluding third molars), it was revealed that out of 682 persons with severe periodontitis, ESI would identify 568, missing 16.7%. Specificity values indicated that ESI would correctly identify 27 (90%) of 30 persons without severe periodontitis. The positive predictive value was 99.4%, as only three persons identified through ESI as affected by severe periodontitis ($n=571$) would not have that level of disease. The negative pre-

dictive value was only 19.1%, however, as 114 (80.9%) persons would be incorrectly considered not to have severe periodontitis. ESI measurements have had significant correspondence with actual increments in severity ($r=0.82$, $P=0.01$) and extent ($r=0.88$, $P=0.01$) values (Figs. 18.15 and 18.16). In Fig. 18.17 are presented variations in apparent prevalence and predictive values derived from ESI measurements when true prevalence is modified. When the latter is 100%, apparent prevalence is only 83%; the positive predictive value tends to increase, and the negative predictive value tends to decrease. Generally speaking, ESI underestimates true prevalence of severe periodontitis, leading to inadequate recommendations for further treatment. Accurately identifying subjects with severe periodontitis requires a full-mouth examination. Because the ESI relies on measurements taken on only 28 periodontal sites to estimate the periodontitis status of the entire mouth, the validity and reliability of ESI may be modified by the prevalence of severe periodontal disease and the distribution of disease according to age and operational definitions (Borges-Yáñez et al. 2004).

18.2.7 O'Leary Gingival Periodontal Index (O'Leary 1967)

The next index was the O'Leary gingival periodontal index formulated in 1967, in which the mouth was divided into six sextants, the highest score for each segment either gingival (0–3) or periodontal (4–6) was recorded, and the sum was divided by the number of segments to give the GPI score for the individual. For periodontal score, the clinical attachment level was recorded at the mesial facial line angle. The main advantages of this index were quick appraisal of the health status of each area of the patient's mouth, and the scores were helpful in determining the personal, facility, and equipment needs of patients (O'Leary 1967).

Fig. 18.15 Correlation between the evaluation of severity using 140 sites and the evaluation of 28 sites (Borges-Yáñez et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

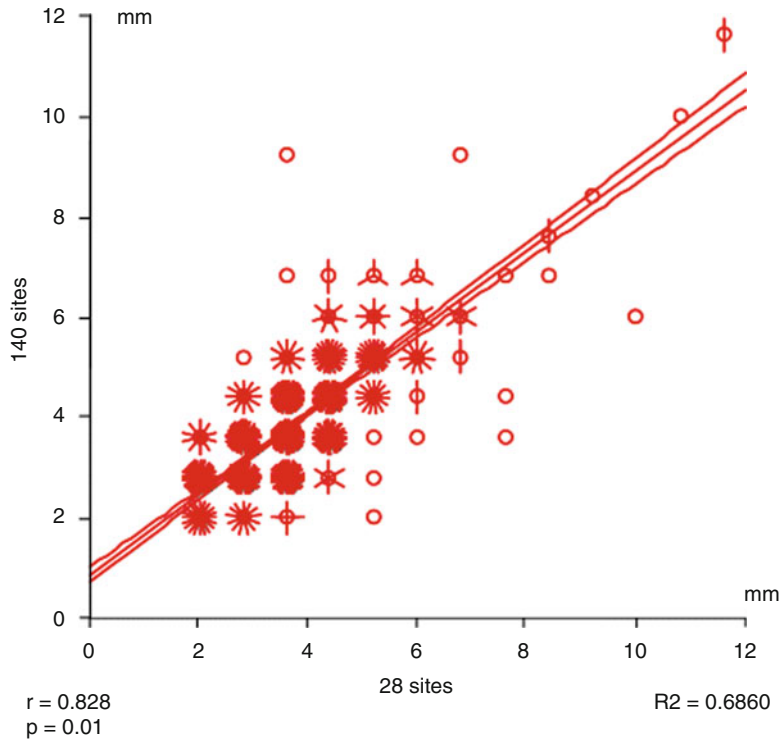
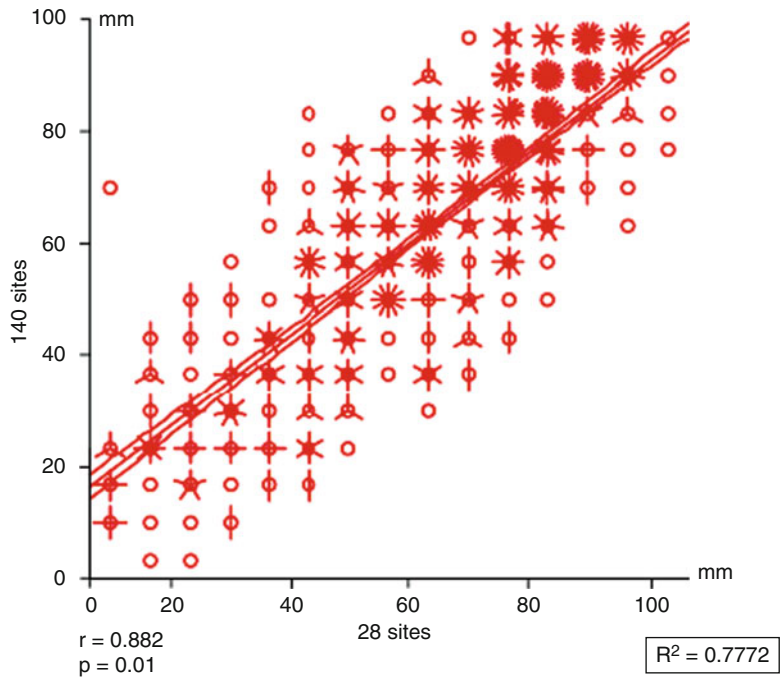


Fig. 18.16 Correlation between the evaluation of extent using 140 sites and the evaluation of 28 sites (Borges-Yáñez et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)



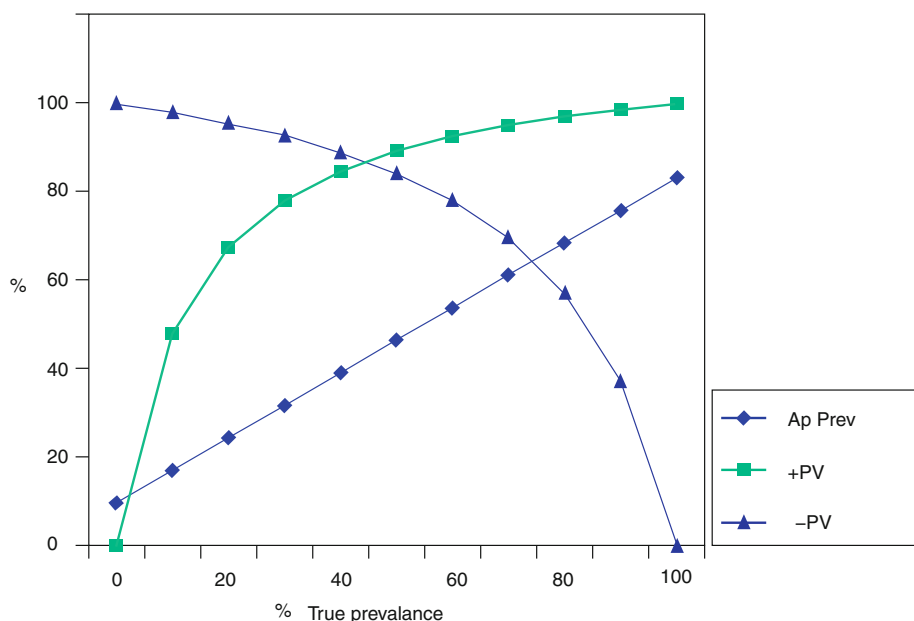


Fig. 18.17 Effect of the true prevalence over the apparent prevalence and predictive values (28 vs. 140 sites) (Borges-Yáñez et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

18.2.8 Navy Periodontal Disease Index (NPDI) (Elliott et al. 1972)

Elliott et al. (1972) tested the severity of periodontal disease using the Navy periodontal disease index applied to six teeth shown to be representative of the total mouth condition (the Ramfjord teeth). The following scores are assessed:

Score 0: gingival tissue normal colour and tightly adapted to the teeth. Tissue has firm consistency. Probing reveals crevicular depth of not over 2 mm, no exudates present.

Score 1: inflammatory changes are present, but these changes do not completely encircle the tooth. These changes may include one or a combination of any colour change from normal gingival colour, loss of normal density and consistency, slight enlargement, and/or blunting of the papilla or the gingival tendency to bleed upon palpation.

Score 2: when the above changes extend to completely encircle the tooth.

Score 6: when the probe reveals pocket depth greater than 2 mm but less than 5 mm.

Score 8: when the probe reveals pocket greater than 5 mm.

The highest combined gingival and pocket score for any one tooth is the patient's NPDI score while NPDI total is the total score for all the six teeth.

Hancock and Wirthlin (1977) examined and scored 98 young adult patients according to the criteria of the Navy periodontal screening examination. The population examined had a 98% incidence of inflammatory periodontal disease. The relationship of the Navy periodontal disease index and the Navy plaque index total scores was determined to +0.55 ($P=0.01$). The relationship of the gingival score of the NPDI and the NPI was found to be +0.75 ($P=0.01$).

The authors recommended the use of the NPDI total as an aid in determining treatment recommendations.

18.2.9 National Institute of Dental and Craniofacial Research (NIDCR) Protocol

In 1985 to 1986, the National Institute of Dental Research, now known as the National Institute of Dental and Craniofacial Research (NIDCR), conducted the National Survey of Oral Health in United States Employed Adults and Seniors. The periodontal assessment used in this survey represented a balance between the newer understandings of periodontal disease epidemiology, the outmoded or limited usefulness of periodontal indices, and the available resources. Instead of implementing an index-based system, probing depth measures were recorded in millimetres, and loss of attachment was calculated. The presence of calculus and bleeding from probing also were recorded. These practices led to the collection of disaggregate site-specific information that permitted researchers to derive a variety of periodontal disease case definitions from the same data source. However, this resource-intensive data collection procedure required the use of a partial-mouth examination to reduce examination time and related costs. Although the US estimates for attachment loss were calculated for the first time, underestimation of disease prevalence became a possibility with the partial-mouth examination procedure (Dye and Thornton-Evans 2007).

In 1988, all of the strengths and limitations of the NIDCR method for periodontal assessment were carried over to the next national health examination survey. NHANES III was conducted by NCHS from 1988 to 1994, and the oral health component of NHANES III was a collaborative effort between NIDCR and NCHS. The study design of NHANES

III was unique compared to prior national health examination surveys. The study was designed to have two nationally representative study periods, or phases, of equal sample size and length. Moreover, the prime goals of NHANES III reflected an expansion of intent beyond providing descriptive information to providing information that would contribute to the understanding of disease aetiology and natural history (Dye and Thornton-Evans 2007).

The health status assessment component of the NHANES III included a detailed interview consisting of demographic, socioeconomic, dietary, and health-related questions. The oral health component of NHANES III included, but was not limited to, the assessment of oral soft tissue lesions, tooth condition, dental caries, removable prosthesis, occlusion and occlusal characteristics, evidence and history of trauma to permanent incisors, and periodontal status. The periodontal examination was performed on subjects 13 years and older. Persons with medical conditions requiring premedication with antibiotic or having certain other conditions were excluded from the periodontal examination. The periodontal examination consisted of the measurement of periodontal supporting tissues including attachment loss, probing pocket depth, and furcation involvement and the assessment of gingival bleeding, dental calculus, and gingival recession. In addition, blood samples were collected for various assays. The periodontal examination was carried out in two randomly selected quadrants, one maxillary and one mandibular. All fully erupted teeth in these two quadrants were assessed, excluding third molars. A maximum of 14 teeth per individual were examined for periodontal parameters (Albandar 2002).

The assessment of the periodontal supporting tissues status was made by clinical measurement of the distance from the cemento-enamel junction (CEJ) to the free gingival margin (FGM) and the distance from the FGM to the bottom of pocket/

sulcus at two sites per tooth, the mesio-buccal and mid-buccal surfaces. The National Institute of Dental and Craniofacial Research (NIDCR) periodontal probe was used. From these two measurements, the probing pocket depth (FGM to bottom of pocket/sulcus), periodontal attachment loss (CEJ to bottom of pocket/sulcus), and gingival recession (CEJ to FGM) were calculated (Albandar 2002).

Assessment of furcation involvement was made on five posterior teeth using explorer #17 (maxillary molars and premolars) and explorer #3 (mandibular molars). Partial furcation involvement (grade I) was scored in sites where the explorer was definitely catching into but did not pass through the furcation. Total furcation involvement (grade II) was used when the explorer could be passed between the roots and through the entire furcation. It is important to note that the NHANES III was the first national survey to assess the periodontal involvement of the furcation area of teeth (Albandar 2002).

18.2.10 Periodontitis Index (Albandar et al. 1999)

Albandar et al. (1999) described a periodontitis index to measure the prevalence and severity of periodontitis in the US population. The periodontitis index classified each person as having either mild, moderate, or advanced periodontitis, or with no periodontitis, based on the number (or percentages) of teeth showing certain thresholds of probing depth and attachment loss. The probing depth was calculated to show the depth of the periodontal pocket that is apical to the cemento-enamel junction around the teeth. Thus, the attachment loss measurement at a given tooth surface was equal to, or greater than, the calculated probing depth measurement (Albandar 2002).

Periodontitis was defined as a disease state in which there was an active destruction of the periodontal supporting tissues as evidenced by the presence of ≥ 3 -mm probing depth and ≥ 3 -mm periodontal attachment loss at the same site. The

individuals were classified according to extent and severity of periodontal disease using the following criteria:

Advanced periodontitis: (1) two or more teeth (or 30% or more of the teeth examined) having ≥ 5 -mm probing depth, (2) four or more teeth (or 60% or more of the teeth examined) having ≥ 4 -mm probing depth, or (3) one or more posterior teeth with grade II furcation involvement.

Moderate periodontitis: (1) one or more teeth with ≥ 5 -mm probing depth, (2) two or more teeth (or 30% or more of the teeth examined) having ≥ 4 -mm probing depth, or (3) one or more posterior teeth with grade I furcation involvement and accompanied with ≥ 3 -mm probing depth.

Mild periodontitis: (1) one or more teeth with ≥ 3 mm-probing depth or (2) one or more posterior teeth with grade I furcation involvement.

No periodontitis: persons with six or more teeth who did not fulfil any of the above criteria were regarded as not having detectable levels of periodontitis.

Each person was assigned a given classification if he/she fulfilled one or more of the criteria of that classification and was given the most advanced classification (Albandar et al. 1999).

The reason for using both the number and percentages of teeth with a given criterion in this classification system is because the NHANES III examined only two randomly selected quadrants (half-mouth), and the use of percentages has the potential to reduce the underestimation due to this partial recording. The index also assessed the extent of furcation involvement of teeth and included that in the assessment of the periodontal status of the person. This is the only periodontitis index that includes a component for the assessment of the furcation involvement of teeth in the appraisal of the severity of periodontitis (Albandar 2002).

This classification system is distinct from other methods in that it assesses multiple parameters of periodontitis, including tissue loss and inflammation, and can be applied to full-mouth data or partial-mouth protocols, which are often used in large national surveys. The index combines measurements of periodontal attachment

loss and probing depth at the same sites in the evaluation of the severity of disease and includes an assessment of the extent of furcation involvement of posterior teeth. It should be noted that the periodontitis index does not regard presence of periodontal attachment loss alone as evidence of periodontitis if the person had no sites with periodontal inflammation or increased probing depth. The periodontitis index classifies the disease into mild, moderate, and advanced periodontitis (Albandar 2011).

When the periodontitis index was applied to the NHANES III (1988–1994) data and the estimates were adjusted for the bias because of the partial-mouth protocol, among the US population aged ≥ 30 years, 30.5% had mild, 13.3% had moderate, and 4.3% had advanced periodontitis. The total adjusted prevalence percentage of subjects with periodontitis was 48.2% (Albandar 2011).

18.2.11 Periodontal Inflamed Surface Area

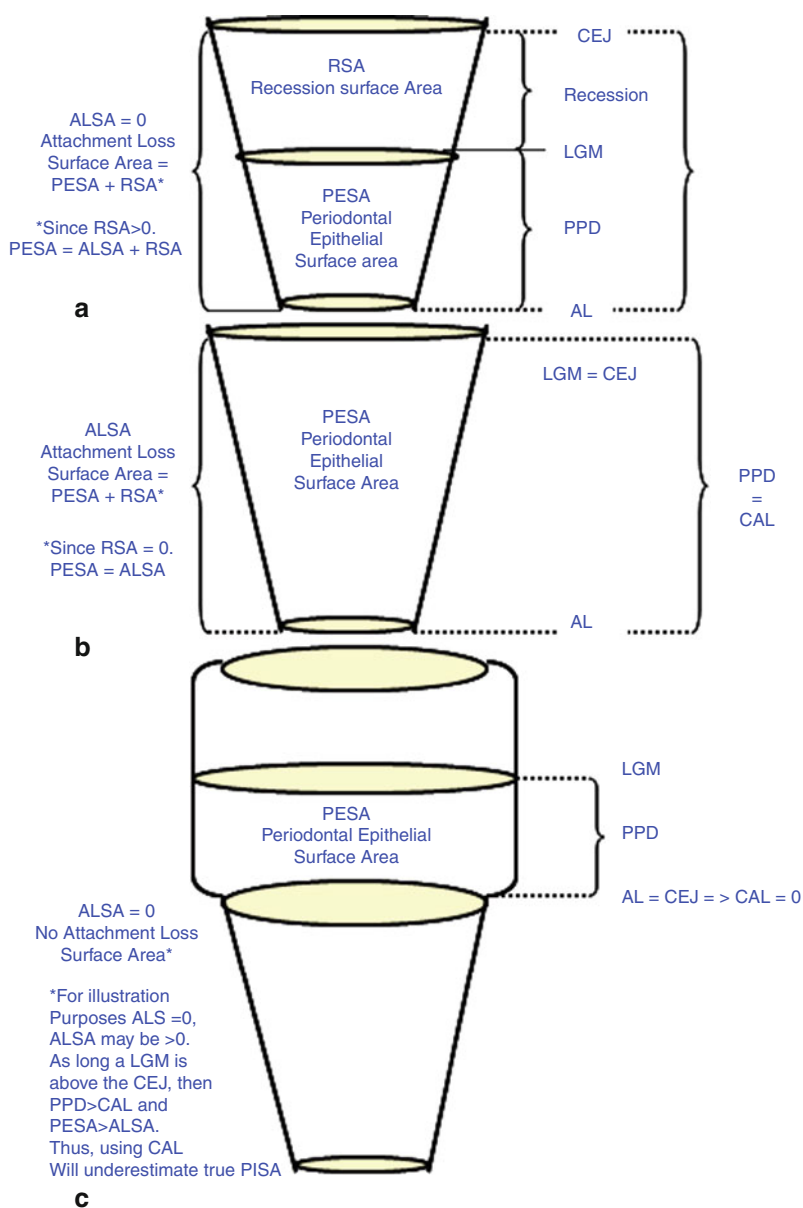
Because no gold standard for periodontitis as a risk factor for other diseases exists, a list of demands was assembled for the construction of a new classification of periodontitis. The first and foremost demand was that the new classification should adequately quantify the amount of inflamed periodontal tissue. Second, the classification should be easy to use and broadly applicable. This means that the classification should make use of clinical measurements commonly used to establish periodontitis, that is, clinical attachment level (CAL) and recession and bleeding on probing (BOP) measurements (Nesse et al. 2008).

Hujoel et al. (2001) related linear periodontal probing measurements to the **dentogingival surface area (DGES)**, also named **attachment loss surface area (ALSA)** (Nesse et al. 2008). The DGES comprises both the sulcular and junctional epithelium, present in health, as well as any intervening pocket epithelium present in periodontitis. Approximate formulas to estimate the dentogingival epithelial surface area

(DGES) were derived from a meta-analysis of the available studies on root surface areas (Hujoel 1994), typical root lengths (Kraus et al. 1969), and a study that related linear loss to surface area loss (Despeignes 1979). Applications of these formulas suggested that the mean DGES among individuals with no periodontitis was 5 cm². Among individuals with periodontitis, the mean DGES in populations ranged from 8 to 20 cm². The maximum DGES observed among the 5,184 individuals in the three populations with periodontitis was 44 cm². The increases in DGES associated with periodontitis ranged somewhere between the ventral surface area of an average adult fingertip (3 cm²) and the ventral surface of one to two average adult fingers (15 cm²) (Hujoel et al. 2001).

The ALSA (DGES) quantifies the root surface area that has become exposed due to attachment loss. However, the ALSA cannot be used to quantify the amount of inflamed periodontal tissue. ALSA does not quantify the periodontal epithelial surface area (PESA) because CAL instead of PPD measurements are used to calculate ALSA (Fig. 18.18). To calculate the PESA, the recession surface area (RSA) has to be subtracted from ALSA (Fig. 18.18a, b). Since $ALSA = PESA + RSA$, it can be deduced that $ALSA - RSA = PESA$. To calculate the PESA, there are three arithmetical possibilities, depending on the location of the gingival margin (LGM) (Nesse et al. 2008):

1. The LGM is below the CEJ so the $RSA > 0$. In this case, $PPD < CAL$ and thus $PESA < ALSA$. Therefore, $PESA = ALSA - RSA$ (Fig. 18.18a).
2. The LGM is exactly at the CEJ. In this case, $PPD = CAL$ and $RSA = 0$. Therefore, $PESA = ALSA$ (Fig. 18.18b).
3. The LGM is above the CEJ. Since $PPD > CAL$, $PESA > ALSA$. Using CAL will lead to an underestimation of PESA. Calculating the PESA is only possible by using PPD instead of CAL, that is, entering PPD as CAL in the formula transforming linear measurements to surface area. This will still lead to an underestimation of PESA (Fig. 18.18c) (Nesse et al. 2008).



List of Abbreviations:

AL = Attachment Level : ALSA = Attachment Loss Surface Area : CAL = Clinical Attachment Level:

CEJ = Cemento-Enamel Junction: LGM = Location of Gingival Margin:

PESA = Periodontal Epithelial Surface Area: PISA = Periodontal Inflamed Surface Area:

PPD = Probing Pocket Depth: RSA = Recession Surface Area

Fig. 18.18 The location of the gingival margin determines how periodontal epithelial surface area is calculated (Nesse et al. 2008. Reprinted with permission from John Wiley & Sons, Inc.)

Thus, PESA accurately quantifies the surface area of pocket epithelium if LGM is at or below the CEJ. However, PESA does still not quantify the surface area of inflamed pocket epithelium. After all, the PESA also includes healthy pocket epithelium. The inflamed part of the PESA on the other hand does theoretically pose an inflammatory burden. To calculate the inflamed part of the PESA, we propose calculating part of the PESA that is affected by BOP (Nesse et al. 2008).

A Microsoft Excel spreadsheet was constructed to facilitate PISA calculation. Spreadsheets are freely available from our website: www.parsprototo.info. All it takes to calculate PISA is filling in CAL, recessions, and BOP on six sites per tooth in this freely downloadable spreadsheet (Nesse et al. 2008).

A dose-response relationship between the amount of inflamed periodontal tissue and HbA1c level, might be indicative for a causal association between periodontitis and type 2 diabetes. The study of Nesse et al. (2009) revealed that the higher the PISA of type 2 diabetics was, the higher their HbA1c levels were. On a group level, an increase of PISA with 333 mm² was associated with a 1.0 percentage point increase of HbA1c, independent of the influence of other factors. PISA quantifies the inflammatory burden posed by periodontitis and can be easily and broadly applied (Nesse et al. 2009).

18.2.12 Dichotomous Periodontal Index

Currently, apart from regular periodontal indices, various dichotomous measurements of periodontitis in the form of presence or absence of pocket or clinical attachment loss against a cut-off point have been assessed (Dhingra and Vandana 2011). The studies are summarized below:

- Dye et al. (2002a, b): Periodontal pocket depth was expressed as a dichotomous variable indicating that at least one dentate site had a pocket depth of ≥ 5 mm. Periodontal probing depths in the range of 5 mm will generally classify an individual as a chronic periodontitis case with moderate destruction (Anonymous 2000).

Both partial and complete furcal bony defect types were grouped to form a dichotomous variable that identified an individual as having either none or at least one dentate site affected by a dental furcation.

- Tezal et al. (2004): Periodontal disease was represented by clinical attachment loss (CAL) and was assessed both as a continuous variable and dichotomized as 1.5 and ≥ 1.5 mm.
- Akhter et al. (2005): Periodontal disease status was assessed using clinical attachment loss (CAL), and the subjects were dichotomized according to mean CAL < 1.5 mm (control group) and ≥ 1.5 mm (diseased group).
- Borrell et al. (2006): The severe periodontitis was defined as a combination of at least 2 interproximal sites with clinical attachment levels of 6 mm or above and at least 1 interproximal site with pocket depths of 5 mm or above. However, these conditions did not need to be present in the same site or tooth. A dichotomous definition was chosen rather than a continuous definition because the former would be more relevant to clinicians and public health professionals.
- Brothwell and Ghiabi (2009): Two periodontal outcome variables were used: dichotomous mean CAL (≤ 2.5 and > 2.5 mm) and the dichotomous severe periodontitis (1 or more sites with > 5 mm CAL).
- Persson et al. (2005): Panoramic radiographs were taken, and radiographic evidence of horizontal alveolar bone loss and vertical bone defects were identified. Bone loss was defined as the distance between the bone level and the cemento-enamel junction ≥ 4.0 mm. Subjects were classified as '0' if they had no radiographic evidence of horizontal, vertical, or any other inter-radicular bone loss; '1' if they had evidence of alveolar bone loss $< 25\%$ of bone height; '2' if the extent of alveolar bone loss varied, on average, between 25% and 50% of bone height, and '3' if the extent of alveolar bone loss exceeded 50% of the root length as a generalized pattern of bone loss. In the dichotomous analysis, bone conditions were defined as having no evidence of alveolar bone loss (score '0') while any other condition was assigned a score of '1'.

18.2.13 Genetic Susceptibility Index for Periodontal Disease

Recently, Moustakis et al. (2007) formulated a genetic susceptibility index (GSI) for both single nucleotide polymorphisms (SNPs) and microbial components of periodontal disease. The researchers took 850 records of 675 Caucasian periodontitis and control patients. The records incorporated genotypes of SNPs (sixty two triplets) like CARD15 (caspase recruitment domain-15) and TGFB (transforming growth factor-b), records of seven bacterial species (*Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis*, *Prevotella intermedia*, *Tannerella forsythensis*, *Peptostreptococcus micros*, *Fusobacterium nucleatum*, and *Campylobacter rectus*), and ethnic origin as well as age, gender, smoking status, periodontal status (pocket depth and attachment loss), and severity assessment (valued over a nominal scale: healthy=0, mild periodontitis=1, and severe periodontitis=2). A statistical process of association rule mining (ARM) was used to derive GSI from genotypes of SNPs (Dhingra and Vandana 2011).



Fig. 18.19 (a) Classic plaque disclosure using erythrosine, note the poor gingival condition associated with the sub-optimal plaque control. (b) Plaque disclosure using PlaqueFinder—a two-tone dye showing areas of older plaque (blue) and newer plaque (red) (Pretty et al. 2005. Reprinted with permission from Elsevier)

18.3 Oral Hygiene Indices

Several indices are recognized as reliable for estimating plaque and calculus area. The choice of the index system to be used in a clinical trial must be made in terms of the objectives of the trial, the size of the population, the period of the study, and the type and extent of the change anticipated. All indices estimate plaque quantitatively either in terms of tooth area covered or the thickness of material in the area measured (Fischman 1988).

As plaque is generally colourless, it is usually stained prior to assessment. Plaque, therefore, within this context, can be described as ‘stainable material’ and will generally include pellicle in addition to the bacterial deposits. Common disclosing agents used include erythrosine (E127, red), and this is combined with a blue dye (E133) in the commercial product Plaque Finder (Pro-Dentec, Cambridge, UK). The solution is rinsed for about

10 s and then expectorated, the dye giving a preferential colour disclosing of older and newer plaque deposits (**Fig. 18.19**) (Pretty et al. 2005).

18.3.1 Simplified Oral Hygiene Index (Greene and Vermillion 1964)

The oral hygiene index has two components, the debris index (DI-S) and the calculus index (CI-S). Each of these indexes, in turn, is based on numerical determinations representing the amount of debris or calculus found on six preselected tooth surfaces. The six surfaces examined for the OHI-S are selected from four posterior and two anterior teeth. In the posterior portion of the dentition, the first molar is examined on each side of each arch. The buccal surfaces of the selected upper molars and the lingual surfaces of the selected lower molars are inspected. In the anterior portion of the mouth, the labial surfaces of the upper right and

the lower left central incisors are scored. In the absence of either of these anterior teeth, the central incisor on the opposite side of the midline is substituted. Only fully erupted permanent teeth are scored. Natural teeth with full crown restorations and surfaces reduced in height by caries or trauma are not scored. Instead, an alternate tooth is examined (Greene and Vermillion 1964).

The following criteria are applied to determine the respective debris and calculus score values for each of the six surfaces examined:

Oral debris—The surface area covered by debris is estimated by running the side of an explorer along the tooth surface being examined. The following scoring system is used:

Score 0=no debris or stain present

Score 1=soft debris covering not more than one-third of the tooth surface being examined or the presence of extrinsic stains without debris regardless of surface area covered

Score 2=soft debris covering more than one-third but not more than two-thirds of the exposed tooth surface

Score 3=soft debris covering more than two-thirds of the exposed tooth surface

Add the debris scores for the six surfaces together and divide by 6. This is the simplified debris index (Lavigne 1998).

Oral calculus—An explorer is used to estimate the surface area covered by supragingival calculus and to probe for the subgingival calculus.

Scores are assigned according to the following criteria:

Score 0=no calculus present

Score 1=supragingival calculus covering not more than one-third of the tooth surface being examined

Score 2=supragingival calculus covering more than one-third but not more than two-thirds of the exposed tooth surface or the presence of individual flecks of subgingival calculus around the cervical portion of the tooth

Score 3=supragingival calculus covering more than two-thirds of the exposed tooth surface or a continuous heavy band of subgingival calculus around the cervical portion of the tooth

Add the calculus scores for the six surfaces together and divide by 6. This is the simplified calculus index (Lavigne 1998).

After the scores for debris and calculus are recorded, the index values are calculated. For each individual, the debris scores are totalled and divided by the number of surfaces scored. At least two of the six possible surfaces must have been examined for an individual score to be calculated. A score for a group of individuals is obtained by computing the average of the individual scores. The average individual or group score is known as the simplified debris index (DI-S).

The same methods are used to obtain the calculus score or the simplified calculus index (CI-S).

The average individual or group debris and calculus scores are combined to obtain the simplified oral hygiene index. The CI-S and DI-S values may range from 0 to 3 and the OHI-S values from 0 to 6 (Greene and Vermillion 1964).

18.3.2 Plaque Index (Ramfjord 1959)

The following teeth were selected as indicators of the periodontal condition within the dentition: Maxillary right first molar 16, maxillary left central incisor 21, maxillary left first bicuspid 24, mandibular left first molar 36, mandibular right central incisor 41, and mandibular right first bicuspid 44. I was recommended to record plaque after application of disclosing solution as follows:

Score 0=no plaque present

Score 1=plaque present on some but not on all of the interproximal and gingival surfaces of the tooth

Score 2=plaque present on all interproximal and gingival surfaces but covering less than one-half of the entire clinical crown

Score 3=plaque extending over all interproximal and gingival surfaces but covering more than one-half of the entire clinical crown (Ramfjord 1959)

The author recommended to record only fully erupted teeth while missing teeth should not be substituted for their examination. The extent of plaque cannot be evaluated without the use of the disclosing solution, which should be used as the final step of the procedures of examination (Fig. 18.20). It should be applied to all of the teeth to be examined at the same time. The scores for plaque for each individual tooth examined are added and the sum



Fig. 18.20 A disclosing agent reveals the plaque accumulation

divided by the number of teeth examined to yield the index of plaque (Ramfjord 1959).

The Ramfjord's plaque index was subsequently modified by Schick and Ash (1961) and has been employed in many clinical trials. The modification of the method involves the examination of facial and lingual surfaces of six selected teeth with the scoring of plaque restricted to the gingival half of the interproximal surfaces. The following scoring system is used:

- 0: absence of dental plaque
- 1: dental plaque in the interproximal area or at the gingival margin covering less than one-third of gingival half of the facial or lingual surface
- 2: dental plaque covering more than one-third but less than two-thirds of the gingival half of the facial or lingual surface
- 3: dental plaque covering two-thirds or more of the gingival half of the facial or lingual surface of the tooth

The total score is obtained by dividing by the number of surfaces examined to derive a mean (Pretty et al. 2005).

18.3.3 Plaque Index (Quigley and Hein 1962)

Quigley and Hein (1962) proposed a new plaque index in order to compare the cleansing efficiency of manual and power brushing. Following use of disclosing dye, the examiner made a quantitative estimate of the amount of stained plaque on the

buccal, labial, and lingual surfaces of the teeth as showed below:

0=no plaque present

1=flecks of stain at gingival margin

2=definite line of plaque at gingival margin

3=plaque extending on gingival third of surface

4=plaque extending on two-thirds of surface

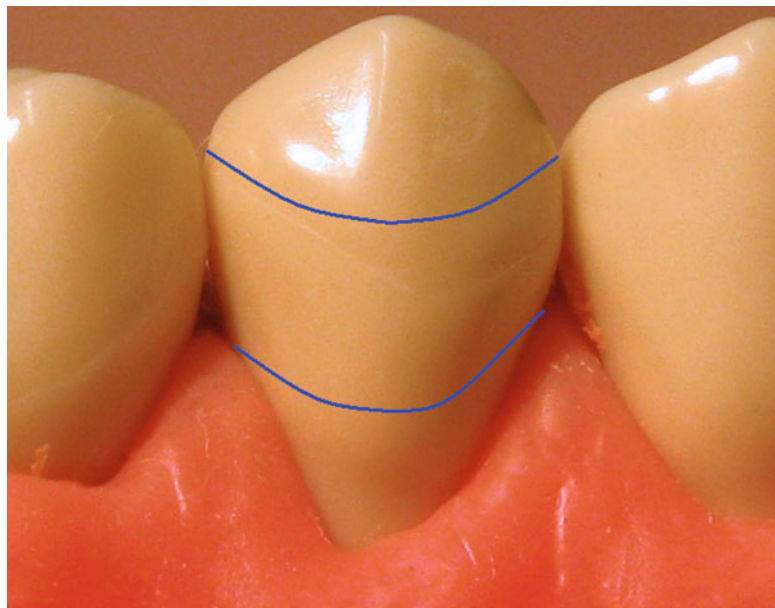
5=plaque extending greater than two-thirds of surface

Evaluation of the data consisted of statistical analysis based on the average amount of plaque per tooth surface per person (Quigley and Hein 1962). This index records the extension of plaque over the tooth surface. Use of disclosing dye means that a score of 1 may be difficult to reach in many patients since a thin film of deposits will result in apposition score. A score of 5 is probably uncommon for teeth in normal alignment since friction from cheeks, lips, tongue, and food tends to keep the coronal parts of the crowns free from plaque. A score of 4 may also be uncommon, at least for people who use a toothbrush on a daily basis (Egelberg 1999).

18.3.4 Plaque Index (Turesky Modified Quigley Hein Plaque Index TQHP) (Turesky et al. 1970)

The plaque index proposed by Quigley and Hein (1962) was subsequently modified by Turesky et al. (1970). Disclosed plaque was scored from 0

Fig. 18.21 The Turesky et al. modified Quigley–Hein plaque index (TQHPI)



to 5 for each facial and lingual non-restored surface only of all the teeth except third molars (**Fig. 18.21**) as follows:

Score 0 = no plaque present

Score 1 = separate flecks of plaque at the cervical margin of the tooth

Score 2 = a thin continuous band of plaque (up to 1 mm) at the cervical margin of the tooth

Score 3 = a band of plaque wider than 1 mm but covering less than one-third of the crown of the tooth

Score 4 = plaque covering at least one-third but less than two-thirds of the crown of the tooth

Score 5 = plaque covering two-thirds or more of the crown of the tooth

An index for the entire mouth was determined by dividing the total all plaque scores by the number surfaces examined.

Modifications of the TQHPI include separating each buccal and lingual aspect into three surfaces (mesial, distal, and mid), using the line angles of the tooth to the contact point bordered by the gingival margin as guidelines for approximal regions, to give a total of six surfaces per tooth (Cugini et al. 2006).

A high degree of consistency within and between the examiners was established (Turesky et al. 1970). This technique of scoring plaque provides a comprehensive method for evaluating

antiplaque agents. This index emphasizes the difference in plaque accumulation in the gingival third of the tooth and tends to overscore the incisal half of the crown, at the expense of the gingival margin (Fischman 1988).

Later, however, investigators have often given separate scores for six aspects of each tooth (mesio-buccal, mid-buccal, disto-buccal, mesio-lingual, mid-lingual, disto-lingual) (Egelberg 1999).

18.3.5 Plaque Index (Silness and Løe 1964)

In the determination of the effect of oral hygiene measures, Silness and Løe (1964) proposed—in the plaque index—determination of the thickness of the plaque at the gingival margin instead of its coronal extension as originally suggested by Greene and Vermillion (1964; Ainamo and Bay 1975).

The teeth which were examined were maxillary right first molar 16, maxillary right lateral incisor 12, maxillary left first bicuspid 24, mandibular left first molar 36, mandibular left lateral incisor 32, and mandibular right first bicuspid 44. Assessment of soft deposits was made according to the plaque index system:

Score 0 = no plaque.

Score 1 = a film of plaque adhering to the free gingival margin and adjacent area of the tooth.

The plaque may be seen in situ only after application of disclosing solution or by using the probe on the tooth surface.

Score 2 = moderate accumulation of soft deposits within the gingival pocket or on the tooth and gingival margin which can be seen with the naked eye.

Score 3 = abundance of soft matter within the gingival pocket and/or on the tooth and gingival margin (Silness and Løe 1964).

Each of the four surfaces of the teeth (buccal, lingual, mesial, and distal) is given a score from 0–3, the plaque index for the area. The scores from the four areas of the tooth are added and divided by four in order to give the **plaque index for the tooth**. The indices for the teeth (incisors, premolars, and molars) may be grouped to designate the index for the group of teeth. By adding the indices for the teeth and dividing by six, the **index for the patient is obtained**. The index for the patient is thus an average score of the number of areas examined. Prior to examination, the gingivae and teeth were dried by a blast of air. No cotton was used in order not to interfere with the soft deposits. For the assessment of plaque, it was found that running an explorer along the surfaces of the teeth both supra- and subgingivally gave better results than the use of disclosing solution and was, therefore, the method of choice (Silness and Løe 1964).

The Silness–Løe Index considers only the thickness of gingival plaque with no consideration of the coronal extension of the biofilm. The index has been criticized because, due to the examination technique, typically, only one examiner can perform the assessment. If this index is to be used, it is recommended that a single, trained examiner be used throughout the trial. The use of the Silness–Løe index in conjunction with other scales, such as the Turesky, has been suggested (Quirynen et al. 1991). However, comparisons between the Silness and Løe index have proved difficult, as the plaque must be disturbed in this methodology. Indeed, a comparison of five indices did not include this popular method (Quirynen et al. 1991; Pretty et al. 2005).

18.3.6 Plaque Control Record (O’Leary et al. 1972)

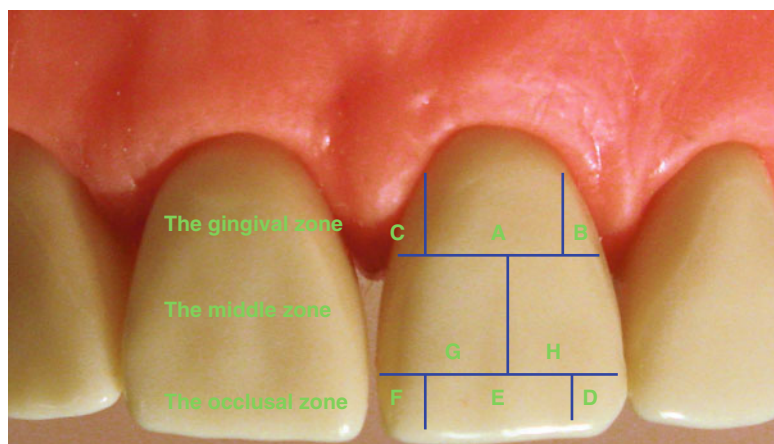
The plaque control record (O’Leary et al. 1972) was developed to give the therapist, hygienist, or dental educator a simple method of recording the presence of plaque of individual tooth surfaces (mesial, distal, facial, lingual). The form also allows the patient to visualize his own progress in learning plaque control. After rinsing the stained solution, each stained surface is examined with an explorer for soft accumulations at the dentogingival junction. When found that they are recorded by making a dash in the appropriate spaces on the record form. No attempt is made to differentiate between varying amounts of plaque on the tooth surfaces. After all the teeth are examined and scored, an index can be derived by dividing the number of plaque-containing surfaces by the total number of available surfaces. The same procedure is carried out at subsequent appointments to determine the patient’s progress in learning and carrying out the prescribed oral hygiene procedures.

One difficulty with this index, and the reason it is used less frequently than the Turesky score, is that it tends to over score the incisal half of crowns at the expense of the gingival margin (Pretty et al. 2005).

18.3.7 Navy Plaque Index (Elliott et al. 1972)

This system of scoring stained plaque is applied on six teeth shown to be representative of the total mouth condition: 16, 21, 24, 36, 41, and 44. The tooth is separated into three major zones, the occlusal, the middle, and the gingival zone. The gingival zone lies apical to an imaginary line connecting the crests of the interdental papillae and roughly parallels the marginal gingival. This area is subdivided into a mesial (C), distal (B), and middle zone (A), with each having a small area, not exceeding 1 mm, adjacent the gingival tissue. The occlusal zone is coronal to the contact area or height of contour (areas F, E, and D). The middle zone extends between the occlusal and gingi-

Fig. 18.22 The Navy scoring system divides the tooth surface into three zones, two of which (middle and gingival) are further subdivided



val zones and is divided into mesial (G) and distal (H) areas (**Fig. 18.22**). By assigning all areas a score of one, more emphasis is placed on plaque adjacent the gingival tissues since the surface area is much smaller. The total score for each tooth is the sum total of all areas of stained plaque on that tooth (Elliott et al. 1972). This index thus requires that the examiner scores as many as 18 different sites per tooth and that the examiner is sufficiently trained to be able to identify and separate the various sites in a consistent manner throughout the dentition. This exacting task may explain why this index has been used less frequently than others (Egelberg 1999).

18.3.8 Modified Navy Plaque Index (The Rustogi Modified Navy Plaque Index RMNPI)

The RMNPI extends the scoring of plaque in approximal (mesial and distal) tooth areas and at the gumline (marginal gingival) region as well as the total tooth. It divides buccal and lingual surfaces into nine areas that are grouped and designated as whole mouth = areas A, B, C, D, E, F, G, H, and I; marginal (gumline) = areas A, B, and C only; and approximal = areas D and F only. Disclosed plaque is scored in each tooth area as present (scored as 1) or absent (scores as 0) and recorded for both buccal and lingual surfaces (**Fig. 18.23**) (Rustogi et al. 1992;

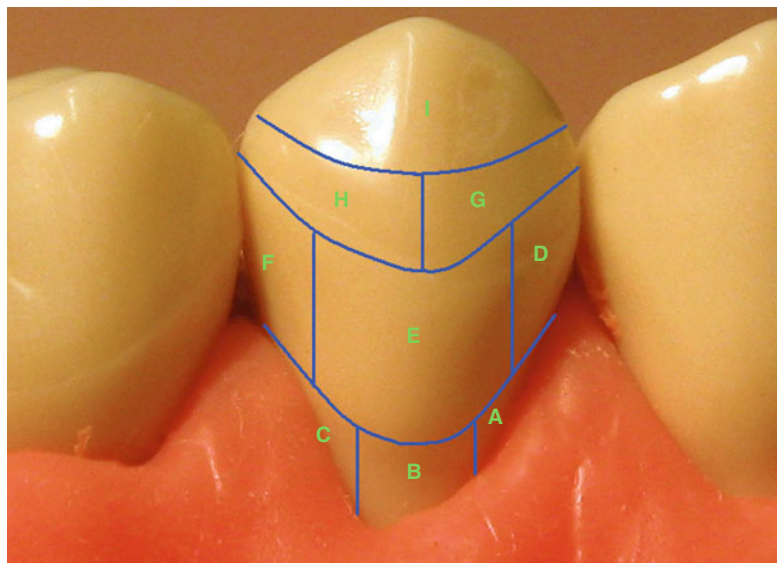
Turesky et al. 1970; Quigley and Hein 1962; Cugini et al. 2006).

18.3.9 No.% of Sites with Plaque

Presence or absence of plaque at the gingival margin is recorded. Plaque scores are expressed as % of examined sites showing plaque. Records may be taken: following use of a disclosing dye; using the tip of a probe moved along the gingival margin to make the plaque discernible or using the unaided dye (Egelberg 1999).

Quantification type indices require a much greater expenditure of time and effort in gathering the data, and the statistical analysis of data derived from such indices is vastly more complex and difficult to perform reliably than the simpler indices which rely on a binary code (i.e., 0 or 1). Thus, binary indices are not only quick and easy to administer but are also highly compatible with algorithms forming the basis of statistical packages generally in use for the analysis of clinical data (Galgut 1999). Presence/absence scores may be an advantage to graded plaque indices since the degree of subjectivity during scoring is reduced. On the other hand, dichotomized presence/absence scores do not allow determination of degrees of plaque accumulation at the individual site. Some investigators have obtained records using a graded index and then presented their findings both using the graded index and using presence/absence scores (positive scores of the

Fig. 18.23 The Rustogi modified Navy plaque index (RMNPI)



graded index versus zero scores; % sites with positive scores of examined sites) (Egelberg 1999).

18.3.10 Approximal Plaque Index (API)

After being stained with erythrosine, a periodontal probe was guided through the approximal spaces of the first and third quadrants from the oral aspect and of the second and fourth quadrants from the buccal aspect. The presence of plaque remnants on the probe was registered. A maximum of 28 measuring points were examined. The number of positive findings was divided by the number of measurements to yield a percent plaque index (Lange 1975). In 20 patients, the inter-rater agreement between the examiner and a senior dentist was assessed. The reliability was $r=0.92$ (Klages et al. 2005).

18.3.11 New Method of Plaque Scoring (NMPS) (Dababneh et al. 2002a)

A new method for plaque scoring (NMPS) was proposed by Dababneh et al. (2002a). According to the NMPS, the visible facial or lingual surface of the tooth is divided horizontally into two major

zones—the gingival one-third (zone A) and the remaining coronal two-thirds of the surface, T, which is further subdivided into three vertical thirds: mesial (zone B), distal (zone C), and middle (zone D). Each of zones A, B, and C is given a score ranging from 0 to 3 depending on subjective evaluation of the proportional area, in thirds, of disclosed plaque on the relevant zone, that is:

Score 0 = no plaque

Score 1 = up to one-third coverage

Score 2 = more than one-third and up to two-thirds coverage

Score 3 = more than two-thirds coverage

The middle third zone, D, is scored on the basis of presence or absence of stained plaque as 1 or 0, respectively. This gives a score ranging from 0 to 10 per buccal or lingual surface (Dababneh et al. 2002a).

The area of interest for the proposed index is the supragingival plaque on the visible smooth (facial/lingual) tooth surfaces, which can include portions of the mesial or distal surfaces, depending on the alignment of the proximal teeth: The alignment of the teeth is often not perfect, and it is possible to find some tilted or crowded teeth within many dental arches. This area may also extend to include half of the proximal surface in the presence of the spaced teeth, missing teeth adjacent to the tooth to be scored, and at the distal surface of the last tooth in the dental arch. A score of 0–3 is given to zones A, B, and C where 0 = no

Fig. 18.24 The criteria for dividing the tooth surface according to the NMPS. Zone A, gingival third; zone B, mesial third; zone C, distal third; zone D, middle third. (Dababneh et al. 2002b)

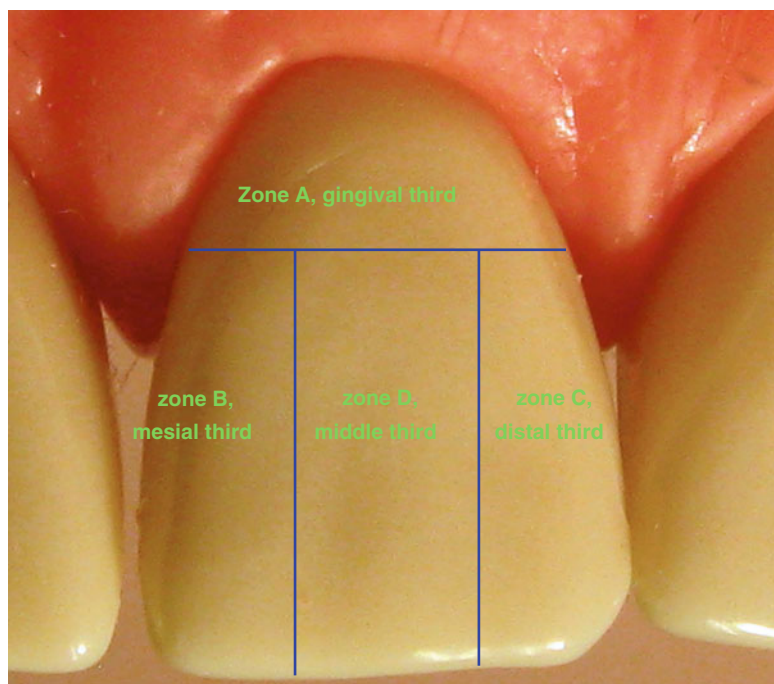
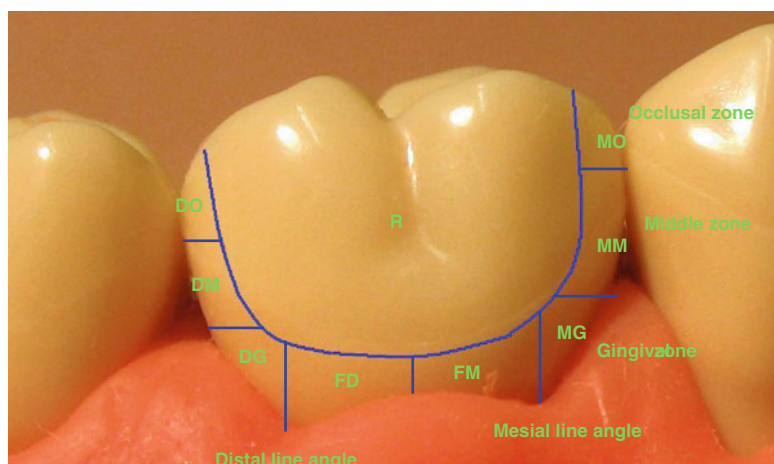


Fig. 18.25 Distal mesial plaque index



plaque, 1=up to one-third coverage, 2=more than one-third and up to two-thirds coverage, and 3=more than two-thirds coverage. A score of 0 or 1 is given to zone D for the presence or absence of plaque (**Fig. 18.24**) (Dababneh et al. 2002a). The aim of the NMPS was to emphasize the plaque scoring at the gingival and proximal regions of the tooth surface (Dababneh et al. 2002a). A positive correlation was found between the NMPS and the Turesky et al. (1970) modification of the Quigley and Hein (1962) plaque index (TPI) ($r=0.717$). The NMPS was found to have less variability

within and between examiners than TPI: Evaluation of NMPS in a clinical setting would now appear warranted (Dababneh et al. 2002b).

18.3.12 Distal Mesial Plaque Index (DMPI)

During recording, the tooth surface is divided into nine areas (**Fig. 18.25**), and each area with the exception of area R can be scored as follows:

0 = no score

1 = plaque covering one-third of the area

2 = plaque covering two-thirds of the area

3 = plaque covering the entire area, and one point is added if plaque present on area R (Fischman et al. 1987; Dababneh et al. 2002a)

18.3.13 The Proximal Marginal Index (PMI) (Benson et al. 1993)

The tooth surface is divided into three unequal segments. The three designated segments are distal proximal, marginal, and mesial proximal. For the distal and mesial areas, the surface scores are defined as follows: from the line angle towards the interdental area, under, and above the contact area as far as visible. The marginal or middle segment is scored from the mesial line angle to the distal line angle, visualizing the area extending incisally/occlusally 3 mm from the marginal gingiva. Plaque in each segment is scored using the criteria described by Turesky et al. modification (1970) of the Quigley and Hein index (1962). Maximum plaque score per surface is 15 (Fig. 18.26) (Benson et al. 1993).

18.3.14 The Axial Plaque Extension Index (APEI) (Matthijs et al. 2001)

The axial plaque extension index (APEI) was recorded with a periodontal probe (PCP UNC 15) to measure the height of the accumulated plaque at the distal and mesial line angles and at the mid-surface of the buccal and lingual surfaces. The probe was held parallel to the long axis of the tooth (Fig. 18.27). At the line angles, the height was measured, starting from the gingival margin to the contact area with the adjacent tooth. Estimates to the closest 0.5 mm were recorded (Matthijs et al. 2001). The APEI scores of the three lingual sites had a substantially lower intra-examiner correlation coefficient than those for the buccal sites, which might be explained by the more difficult accessibility to these surfaces (Table 18.3).

18.3.15 The Proximal Plaque Extension Index (PPEI) (Matthijs et al. 2001)

The proximal plaque extension index (PPEI) was recorded at the mesial and distal part of the buccal and lingual tooth surfaces, starting from the gingival margin. The periodontal probe was kept parallel to an imaginary diagonal line running perpendicularly to the interdental papilla. Estimates to the closest 0.5 mm were recorded (Fig. 18.28) (Matthijs et al. 2001).

18.3.16 Problems with Plaque Indices

There has been much criticism of the use of plaque indices, relating particular to their resolution. For example, if one were using the Quigley Hein index, it would be possible for a tooth's plaque level to decrease by 50% and yet would still be scored the same. Other difficulties lie with the subjective nature of the indices and the need for examiner training, often increasing the cost of clinical trials, as does the need for a clinician to conduct the examination. The ability to ensure that, within any one trial, all examiners are calibrated does not necessarily confer reliability to trials conducted within other centres or at differing times within the same centre (Pretty et al. 2005).

The use of indices for plaque quantification lacks the precision, objectivity, sensitivity, specificity, and reliability that the highest level of clinical trial design requires. It is correct to say that examiners have been extensively calibrated to ensure intra and inter-examiner agreements, but this is costly and time-consuming. Tests that enable a more sensitive detection of plaque, with a higher discriminatory power, may enable both the number of participants in and amount of time required for clinical trials to be reduced. This is of obvious benefit to clinical researchers and product manufacturers. Indeed, not only are researchers studying small changes in plaque, they are doing so within very small plaque amounts. The ability to use interval rather than ordinal data also lends statistical strength to trials. Planimetric methods may be a solution to this problem (Pretty et al. 2005).

Fig. 18.26 Proximal marginal index (Benson et al. 1993)

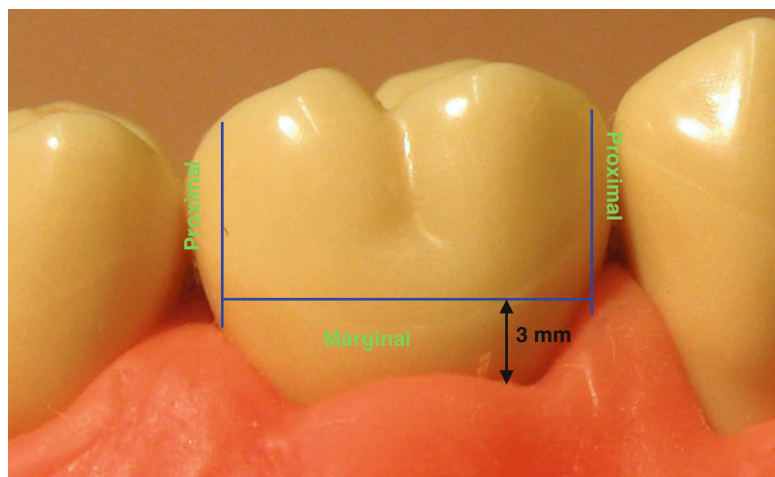
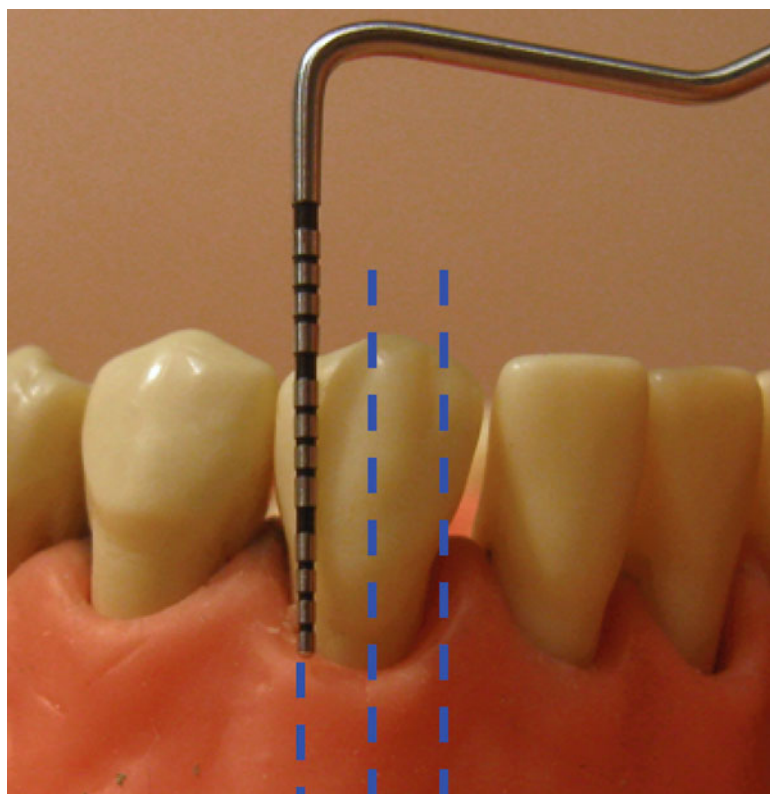


Fig. 18.27 Axial plaque extension index. The index is recorded by holding a periodontal probe parallel to the long axis of the tooth (Matthijs et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)



18.3.17 The Plaque Area (PLA) Measurements

The most frequently used plaque indices only score presence of plaque related to some areas of interest (for instance, the gingival third of the tooth). Although most other plaque indices use

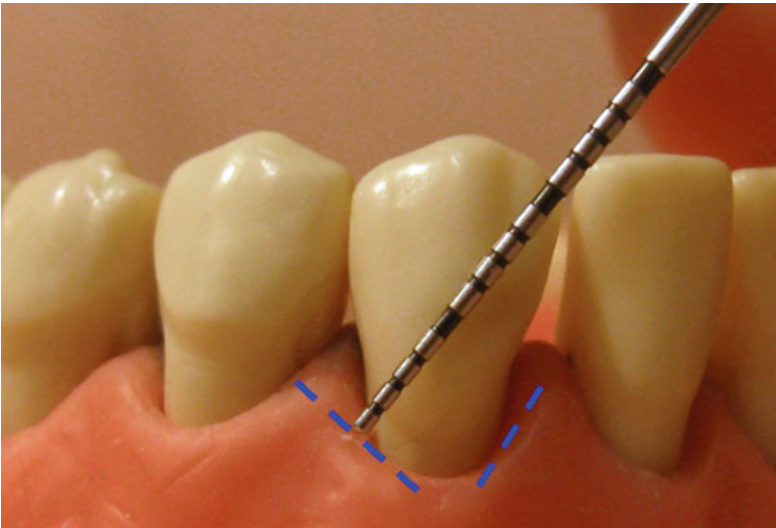
simple categorical definitions, they are not strictly quantitative. For example, dividing the tooth surface into segments only represents a quantized measurement because the score only accounts for how much of the segment is covered by plaque. This means that the same score on teeth of equal size may be due to different amounts of plaque.

Table 18.3 Reproducibility of the visual plaque index and the modified navy plaque index for the buccal and lingual surfaces (Matthijs et al. 2001) (Reprinted with permission from John Wiley & Sons, Inc.)

	Visual plaque index		Modified navy plaque index	
	Buccal	Lingual	Buccal	Lingual
<i>rs</i> (Spearman)	0.90	0.85	0.89	0.85
% agreement	78	89	88	73
<i>k</i>	0.70	0.75	0.81	0.62
MME (%)	3.8	2.8	0.5	1.3

MME mean measurement error

Fig. 18.28 Proximal plaque extension index. The index is recorded by holding a periodontal probe parallel to an imaginary *diagonal line*, perpendicular to the interdental papilla (Matthijs et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)



The opposite may also be true: Using tooth segments to obtain scores effectively disregards tooth size, meaning that the same amount of plaque may yield different scores on different sized teeth (typically, central and lateral incisors) (Carter et al. 2004).

It was shown that plaque area tracings can be rescored into scores for an area-related plaque index. The repeatability of examiners was good for the rescoring method and with the original index. This supports the idea that plaque area recorded graphically or photographically provides a permanent record of plaque distribution on teeth that can be reassessed at a later date and away from the pressures of the clinic environments. The recording of data in this way would allow studies to be compared when initial indices recorded differed between studies (Renton-Harper et al. 1999).

Superiority of the index compared to the area was shown by Addy et al. (1999) who evaluated

the results of 15 plaque regrowth studies performed on healthy volunteers. In two-thirds of these studies, the plaque index showed a greater discriminatory power compared to plaque area measurements (Shaw and Murray stain index (1977) modified by Addy et al. (1983)), while the other one-third showed a lower discriminatory power. Planimetric evaluation in these studies was performed on the upper and lower incisors, canines, and premolars. The authors discussed the reduced number of selected teeth compared to the full-mouth index approach in order to explain the lower discriminatory power of the plaque area assessment in most of the studies. In contrast, Quirynen et al. (1991) found that planimetric area measurements were superior to five other plaque indices. These authors included three randomly selected teeth per subject while comparing undisturbed plaque growth over a period of 96 h on patients who had healthy and inflamed gingival (Lorenz et al. 2009b).

However, it was shown that plaque area tracings can be rescored into scores for an area-related plaque index. The repeatability of examiners was good for the rescoring method and with the original index. This supports the idea that plaque area recorded graphically or photographically provides a permanent record of plaque distribution on teeth that can be reassessed at a later date and away from the pressures of the clinic environments. The recording of data in this way would allow studies to be compared when initial indices recorded differed between studies (Renton-Harper et al. 1999). Similar results were reported by Smith et al. (2006) who showed that a computerized method of image analysis gave comparable trial conclusions to the alternative methods of plaque quantification used; Smith et al. (2001) reported measurement of dental plaque area on the labial surface of anterior teeth and Smith et al. (2004) modification for the lingual surface of anterior teeth. The planimetric analysis has also the advantage of being a permanent record that can repeatedly be re-evaluated (Lorenz et al. 2009b). Other advantage is being able to accurately quantify the oral surface covered by plaque, rather than using indices. For instance, it is possible to assess in detail how one cleaning technique performs against another since such performance is a quantitative variable. Objective detection of plaque also allows accurate spatial localization of plaque in relation to the oral structures; the most obvious is the extent of plaque in close contact with the cervical margin and at the margins of restorations (Carter et al. 2004). However, in terms of costs and time effort, the plaque area assessment particular on multiple teeth seems to be disadvantageous compared to an index when applied in clinical studies on mouth rinses (Lorenz et al. 2009b).

18.3.17.1 Computer-Based Analysis of Dental Plaque

The planimetric indices can include several numbers of teeth (Arweiler et al. 2001a, b; Quirynen and van Steenberghe 1989). For example Arweiler et al. (2002a, b) measured plaque area at four upper and four lower incisors while Lorenz et al. (2009b) selected only one tooth (the

upper right central incisor) for area estimations. The percentage of the plaque covered surface of these teeth was recorded after staining with erythrosine, digital standardized photographing, and computer-based calculation. Orthoradial photographs were taken in order to assess the tooth surface in its full extent. The stained buccal surface was highlighted on the photograph using the drawing tool of Adobe Photoshop 7.0 software, and then the number of pixels within this area was calculated. In addition, the circumference of the whole tooth surface was also highlighted and numbers of pixels within this area calculated. The relation between the plaque covered area (number of pixels) to the total vestibular tooth surface (number of pixels) gave the percentage of existing plaque (Lorenz et al. 2009a, b; Renton-Harper et al. 1999).

Most of the planimetric plaque evaluation methods were based on photographic images of the buccal surfaces of anterior teeth (Shaloub and Addy 2000), and only few studies dealt with all tooth surfaces, which were photographed using mirrors (Söder et al. 1993, 1995). However, none of the studies focused on the tooth surfaces difficult to reach. Recently, Staudt et al. (2001) introduced a computerized planimetric method, using an intraoral CCD camera in combination with a newly developed positioner, which allowed standardized and reproducible images of lingual surfaces of mandibular teeth in aim to compare the effectiveness of three toothbrushes with different brush head designs on the lingual surfaces of mandibular premolars and molars.

For the computer-based planimetric plaque analysis, a modification of the plaque percent index (PPI, Söder et al. 1993) was calculated using a software developed especially for this purpose (med3D, Heidelberg, Germany). In a first step, each digital image was filtered by removing eight specific colours, which were to be used in a later stage for counting. In our case, only two of the specific colours were used. These specific colours were replaced by unperceptible different ones (e.g. black, consisting of the values for red=000, green=000, blue=000 became R=001, G=001, B=001). By means of an already existing software (Corel Photo-Paint, Corel

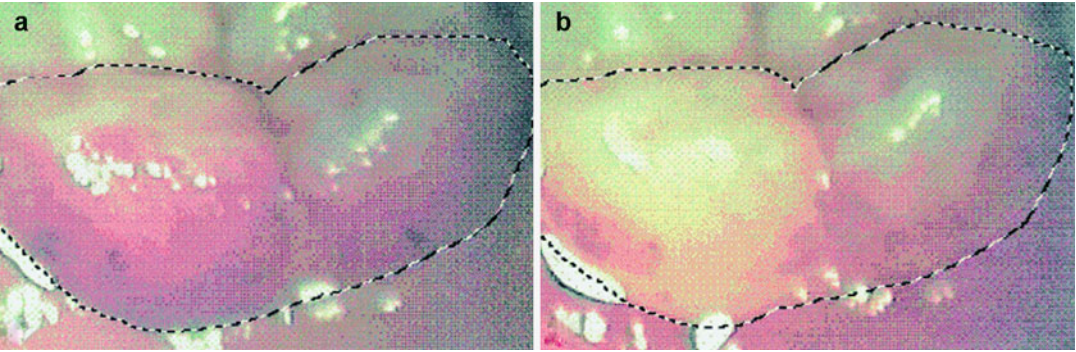


Fig. 18.29 Tooth 37 with stained plaque before (a) and after (b) toothbrushing. Entire tooth-area is outlined by a mask (Staudt et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)

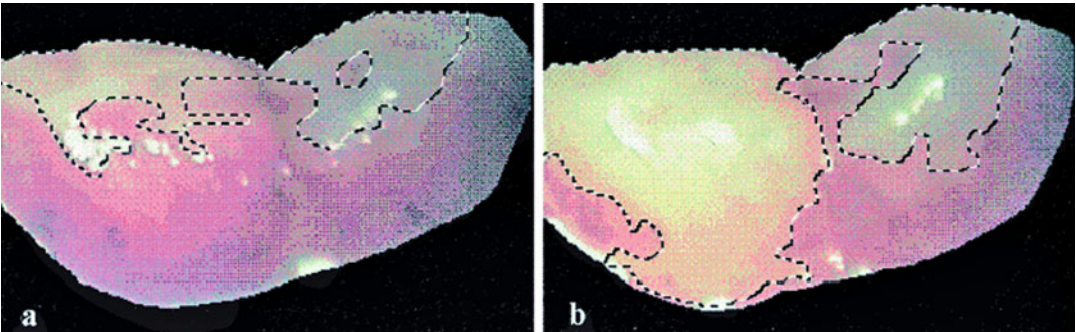


Fig. 18.30 Tooth 37 with stained plaque before (a) and after (b) toothbrushing. Background colours are replaced with black. Tooth-area without plaque is outlined by a second mask (Staudt et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)

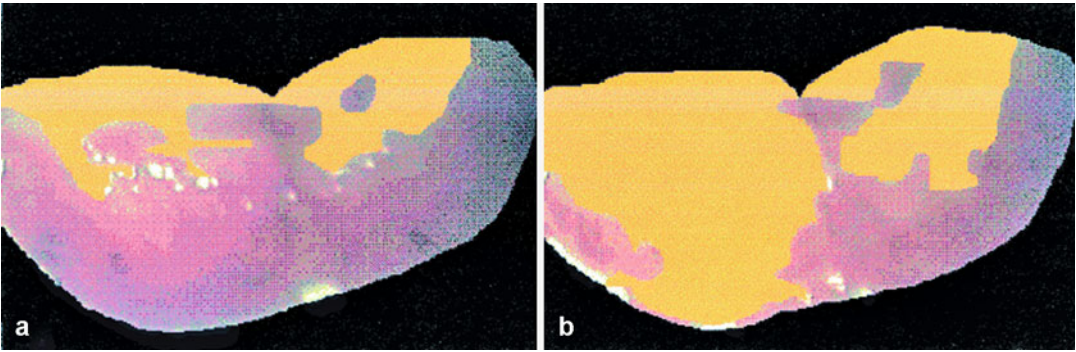


Fig. 18.31 Tooth 37 with stained plaque before (a) and after (b) toothbrushing. The colours in the second mask are replaced with yellow. Plaque-area remains as surface with natural colours (Staudt et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)

Corp., Ottawa, Canada), the entire tooth area (Fig. 18.29) and the tooth area without plaque were outlined by a mask (Fig. 18.30). These masks were then filled each with one of the specific colours to determine tooth area (= entire

picture—background [filled with black]) and plaque area (= entire picture—background [black]—tooth area without plaque [yellow]) (Fig. 18.31). Again, the images were analyzed by the newly developed software, this time by count-

Fig. 18.32 Masking of 12 front teeth (6 maxillary, 6 mandibular) (Klukowska et al. 2011. Reprinted with permission from Elsevier)

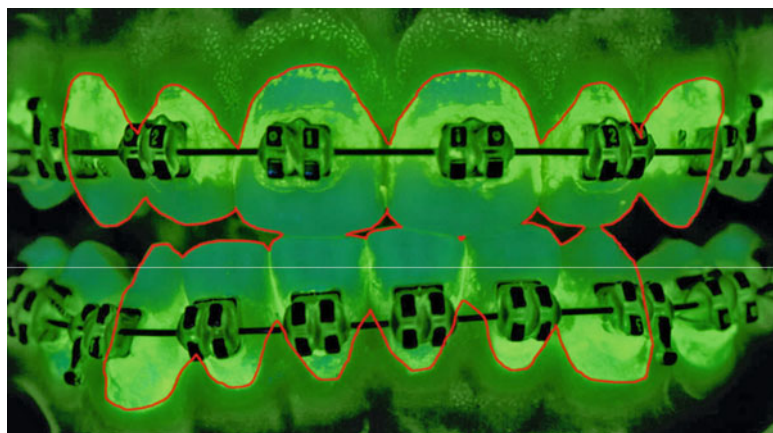
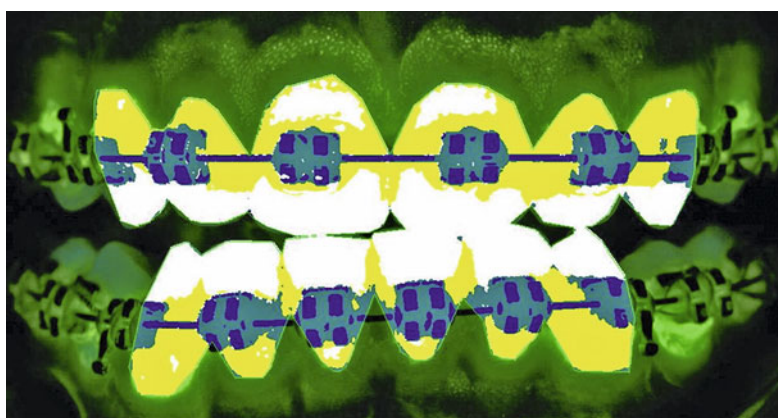


Fig. 18.33 Discrimination rule for orthodontic patients, including plaque on tooth (yellow), teeth (white), plaque on the bracket (light blue), and brackets (dark blue) (Klukowska et al. 2011. Reprinted with permission from Elsevier)



ing for each of the specific colours the number of pixels contained in the full image. Plaque area was then expressed as percentage of tooth area (plaque percent index) (Staudt et al. 2001).

18.3.17.2 Fluorescein Disclosing and Digital Plaque Image Analysis (DPIA)

A fluorescein plaque-disclosing agent provides the solution to problems associated with using Red Coat. Fluorescein, FD&C yellow No. 8, is a UV fluorescent dye that absorbs into dental plaque in a fashion very similar to Red Coat, not surprising considering the chemical structural similarity between Red Coat and fluorescein. Under long-wave UV light, fluorescein produces a disclosed plaque on teeth that is significantly different from the colour of gingival tissue and gingival plaque. Plaque on teeth fluoresces yellow, plaque on gingiva fluoresces green, teeth

fluoresces blue, and gingiva are nearly black. These unique properties of fluorescein make it an excellent candidate for automatic determination of plaque coverage by computer-aided imaging using colour as the classifying feature. The five distinct classes under UV light, plaque on the teeth, plaque on the gingiva, teeth, gums, and lip retractors, are separated by at least 50 colour values (255 maximum) in red, green, and blue (RGB) colour space. Captured images are classified with Optimas® macros. The image pixels are classified using the least squared distance model. The results of each analysis are written to an ascii text file and include subject ID, visit ID, analysis date, number of pixels of teeth, number of pixels of gingival plaque, number of pixels of plaque on teeth, and percent plaque on teeth. **Figure 18.32** is an actual image, and **Fig. 18.33** shows a generated classification image using the plaque-imaging system (Sagel et al. 2000).

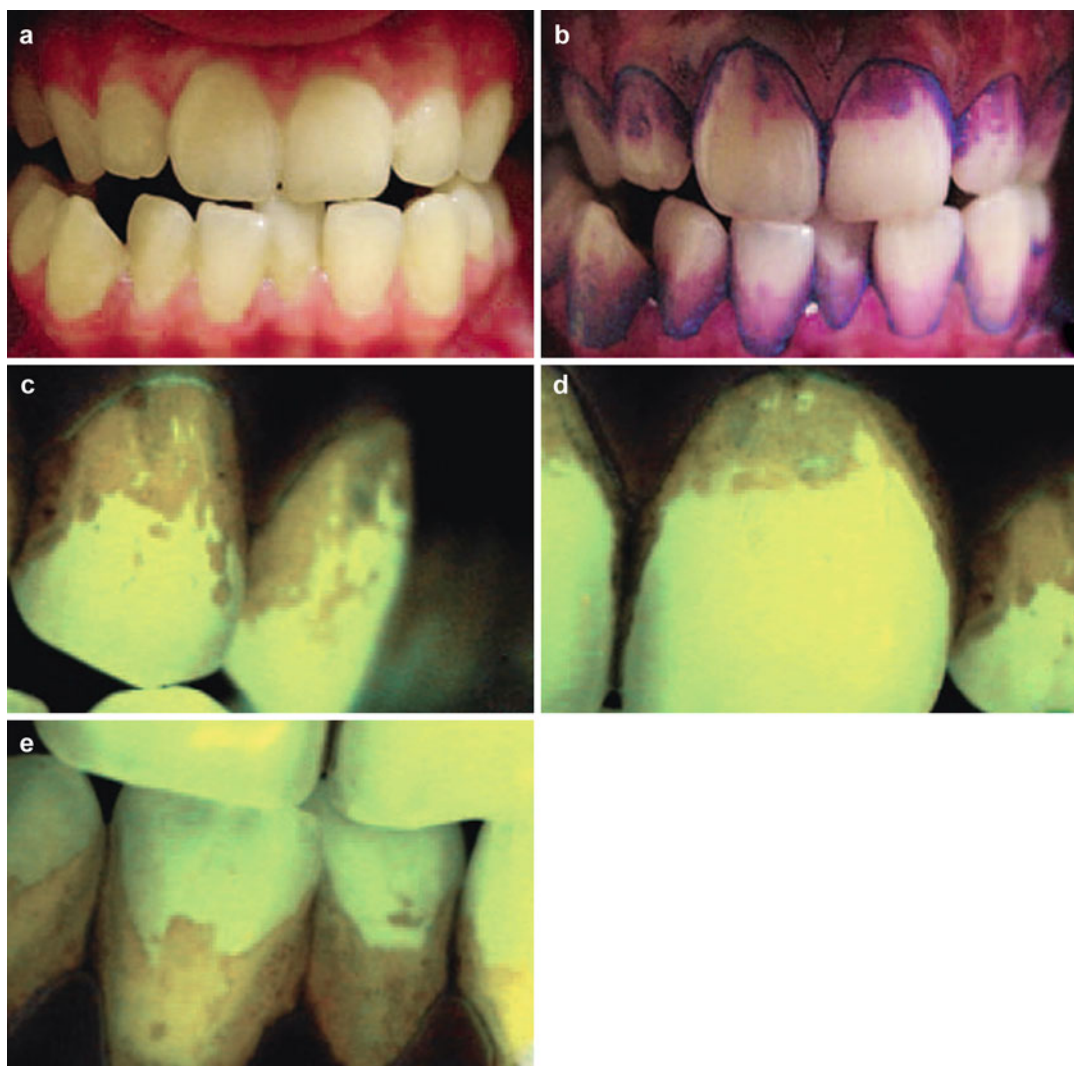


Fig. 18.34 (a) Photograph of teeth prior to disclosure with plaque finder, (b) photograph of teeth following plaque disclosure, and (c, d and e) examples of quantitative light-

induced fluorescence (QLF) images of disclosed dental plaque (Pretty et al. 2004. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

18.3.17.3 Plaque Detection with Quantitative Light-Induced Fluorescence (QLF)

The advent of quantitative light-induced fluorescence (QLF) to detect early carious lesions has offered a novel way of obtaining images suitable for planimetric analysis. Under QLF conditions, plaque is visible as an orange area, the teeth are green in colour, and the gingivae are black or brown. These elements are very discrete, and, therefore, it is proposed that

such images would be highly amenable to planimetric analysis as the image analysis software can easily distinguish between these colours (Fig. 18.34). Digital photographs and QLF images were taken of the anterior teeth (canine to canine, Mn, Mx) for planimetric plaque analysis. A headrest assembly was used to ensure that angulations for each of the images was standardized (Fig. 18.35). Planimetric analysis was performed using an image analysis package (Scion Image, Scion Corporation, USA) by



Fig. 18.35 QLF within the clinical setting with a headrest for reproducible images (Pretty et al. 2004. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)

a single examiner using the images acquired by the QLF device and the conventional digital photographs. QLF images were available for the maxillary and mandibular anterior teeth on a total of 12 surfaces for analysis. Initially, each tooth was selected from the image and subsequently the plaque from that tooth, defined by the orange/red colour removed. Each of these two images (the total tooth image and the plaque only image) was measured at a pixel level. **Figure 18.36** shows an example of this analysis. For the conventional method, the blue-red areas were selected and measured in the same manner. From the pixel measurements, the percentage plaque index was calculated and entered into *SPSS* for later analysis (Pretty et al. 2004). Also, a method for the detection of plaque with QLF but employing no disclosing solution, relying purely on the natural fluorescence of the plaque, is also described. This QLF analysis of plaque has many advantages. The small camera, which can either be handheld or used within a chin-rest system, is easy to use. Images are free from flashlight, distortion, and specular reflections, all of which will affect planimetric analyses

(**Figs. 18.37** and **18.38**). The introduction of a new video repositioning system (VidRep, Inspektor Research Systems, NL) promises to increase the reliability of the device further by ensuring that the angulation of camera to tooth is consistent when monitoring longitudinally or through different legs of a clinical trial. As with all planimetric methods, the QLF plaque assessment will enable very small changes in plaque to be detected and hence will reduce the time taken in clinical trials to determine any significant effects. It is also possible that QLF can be used to assess denture plaque distribution (Pretty et al. 2005).

18.3.17.4 Plaque Quantification Using 3D Coordinate Data

The work of Jovanovski, Yeganeh, Lynch, and others in the use of 3D mapping of tooth structures has been innovatively applied to the quantification of plaque. Using a 'coordinate measuring machine' (CMM), with replicas of the tooth surfaces, 3D images of the surface are acquired using an optical probe (Renishaw OP2 laser probe). The laser is a Class IIb, GaAl-As with a maximum power output

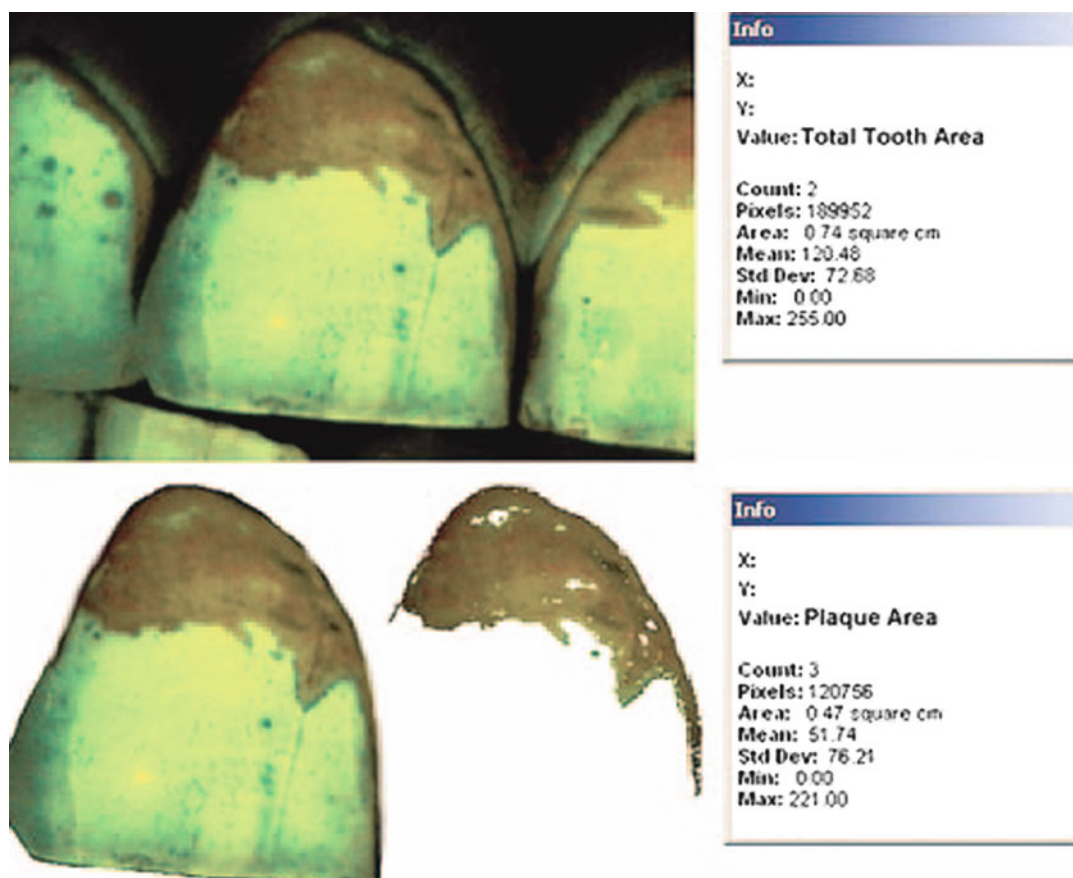


Fig. 18.36 Example of the planimetric measurement technique (Pretty et al. 2004. Reprinted with permission from Nature Publishing Group: Macmillan Publishers Ltd.)



Fig. 18.37 The QLF handpiece with mirror attached. The QLF device can image all areas of the mouth, including lingual and occlusal surfaces (Pretty et al. 2005. Reprinted with permission from Elsevier)

of 5.0 mW. In order to scan oral tissues, the tooth surface of interest must be replicated, as this is an *in vitro* device, whose mode of operation of and

physical characteristics preclude its use *in vivo* (Pretty et al. 2005; Jovanovski and Lynch 2000). Jovanovski describes the use of a polyvinyl siloxane impression material that is placed within a specially designed (for each area of interest) impression tube. The use of this tube enabled repeated impression of the same oral structure, ensuring the same orientation on each occasion (Jovanovski and Lynch 2000).

Yeganeh et al. (1999) quantified the plaque thickness at various levels adjacent to the gingival margin using a three-dimensional method and to correlate these measurements to the plaque index described by Silness and L  e (1967). The Silness and L  e plaque index was scored, and replicas were scanned using a coordinate measuring machine (CMM) and laser scanning probe. A replica was obtained from this surface before

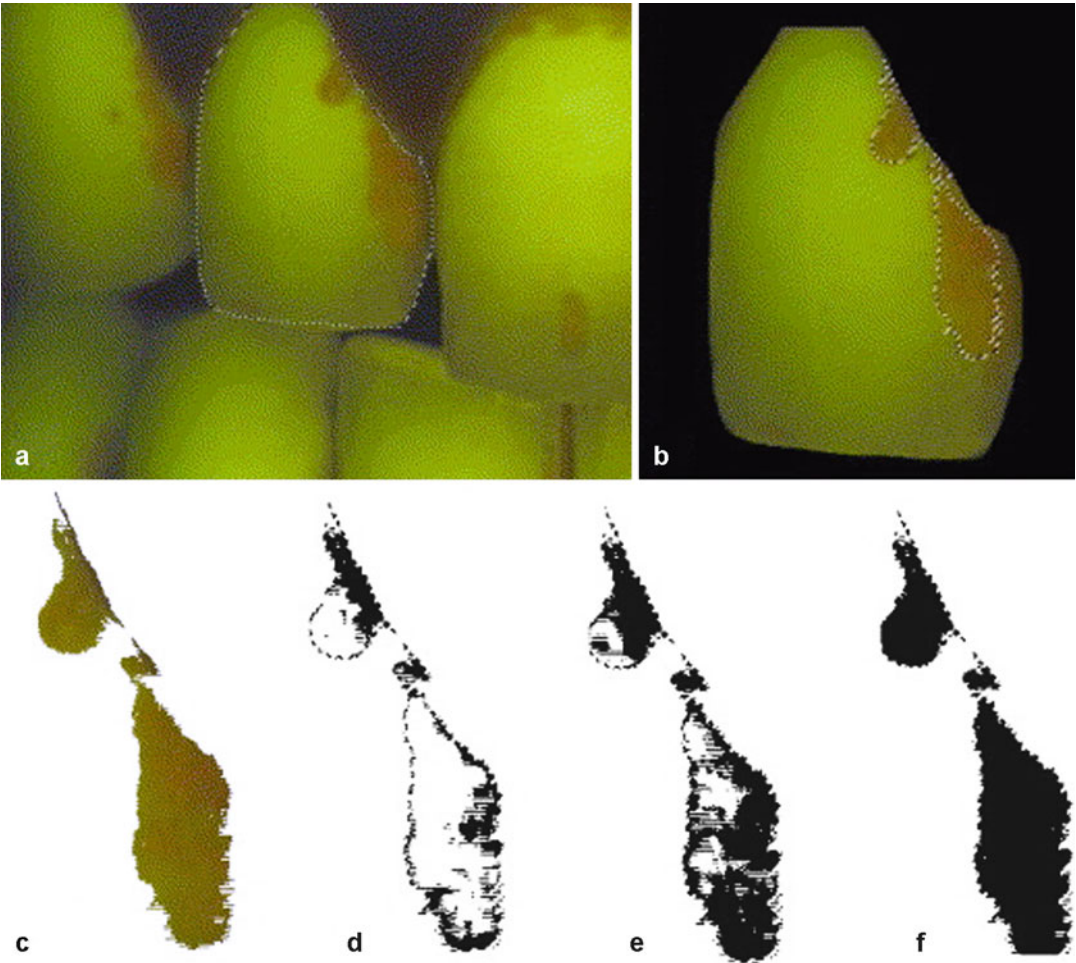


Fig. 18.38 Planimetric analysis of undisclosed plaque. (a) QLF image taken of the surface under study. (b) Individual tooth selected and plaque automatically detected by the image analysis software. (c) The plaque area at a variety of thresholds, which may relate to plaque

density or depth. The percentage plaque index (PPI) in this case is 7.9%, or 6,641 plaque pixels out of a total tooth pixel count of 83,319 (Pretty et al. 2005. Reprinted with permission from Elsevier)

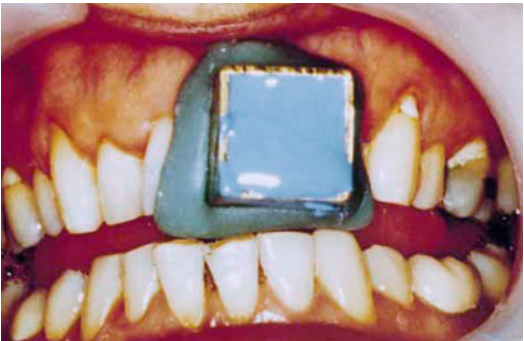


Fig. 18.39 A replica of the site is taken using light-body-addition silicone-impression material (Yeganeh et al. 1999. Reprinted with permission from John Wiley & Sons, Inc.)



Fig. 18.40 Four sequential replicas of plaque on root surface taken to assess repeatability (Yeganeh et al. 1999. Reprinted with permission from John Wiley & Sons, Inc.)

and after toothbrushing (**Figs. 18.39, 18.40, and 18.41**). The mean plaque thickness was measured at the gingival margin and ranged from 0.015 to 0.029 mm, from 0.043 to 0.066 mm, from 0.141 to 0.271 mm, and from 0.336 to 0.560 mm when PI of 0, 1, 2, and 3 were respectively recorded by the clinical examiner, showing a strong linear relationship between the plaque thickness and the plaque index recorded by the same examiner. The mean plaque thickness measured by the system adjacent to the gingival margin increased with higher PI. Plaque adjacent to the gingival margin had a mean thickness of 0.106 ± 0.118 mm (mean $\pm$ SD) whilst mean plaque thickness 250 μ m from the gingival margin was 0.053 ± 0.052 mm (mean $\pm$ SD) (Yeganeh et al. 1999).

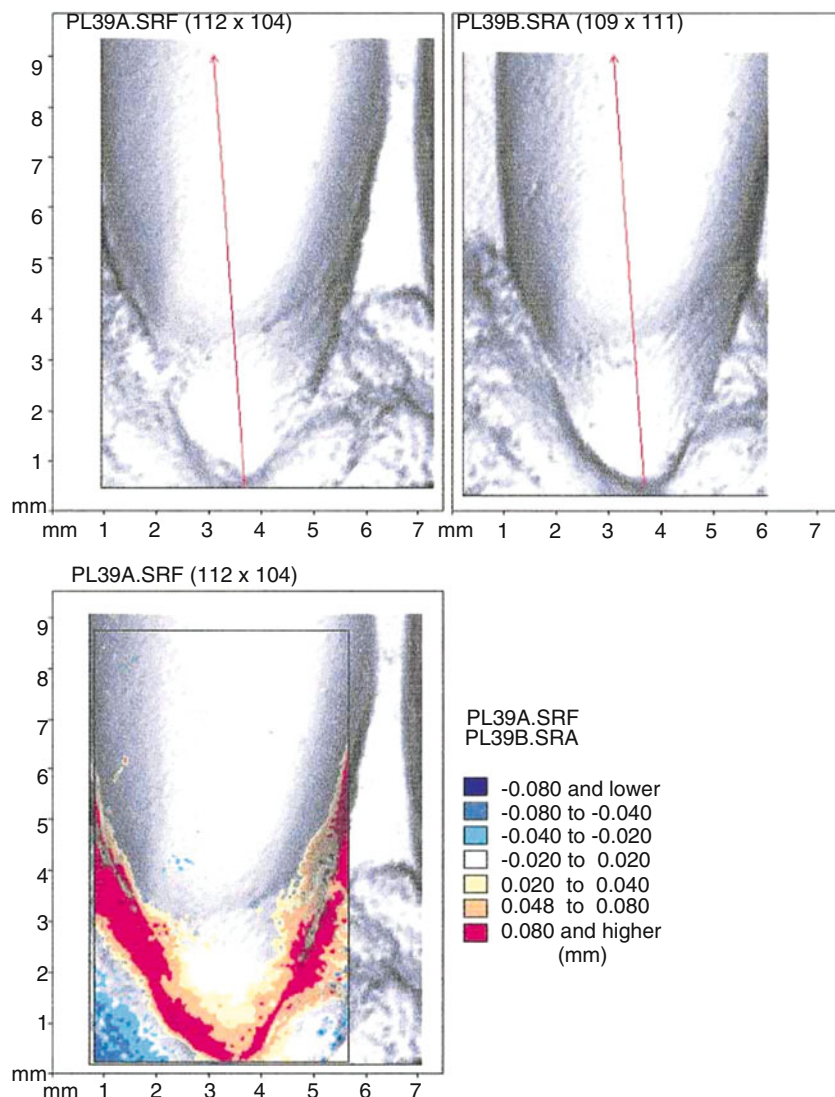
While a very sensitive and valid technique, the laser scanning system suffers from its inability to be used within the mouth. The use of direct impressions of the area of interest, and the ability to scan these impressions, rather than poured stone casts, removes variability that may be caused by shrinkage or other casting errors. However, the use of an impression material may disturb, compress, or remove plaque, possibly allowing for false negatives. The system's upper limit of sensitivity for plaque is determined by the

resolution of the impression material. Nonetheless, the system provides a wealth of information about the oral tissues, for example, it would be possible to measure the effect of a test product on gingival contour. The assessment of the morphological features of the teeth, soft tissues, and plaque during clinical studies offers exciting opportunities in the future (Pretty et al. 2005).

18.3.18 Quantification of the Thickness of Plaque Biofilm

There is no doubt that plaque indices have considerable advantages, most notably in that they are quick to use and, therefore, efficient in the clinical trial situation. There are, however, disadvantages also. First, plaque indices are used to score plaque deposits on the clinical crown and do not assess subgingival deposits. Second, the indices are entirely subjective, particularly at the hard to visualize interproximal sites, and their use in clinical trials necessitates intra and (when there is more than one examiner) inter-examiner calibrations that usually only achieve κ statistics in the range [0.60–0.75]. Expected and standardized differences and power and sample size calculations must allow for this degree of potential subjectivity

Fig. 18.41 Two sequential replicas made immediately before and after removal of plaque. A colour-coded difference map shows the plaque distribution (Yeganeh et al. 1999. Reprinted with permission from John Wiley & Sons, Inc.)



and variability, thus necessitating a need to recruit relatively large subject cohorts leading to time-consuming and expensive clinical trials. Third, plaque indices are categorical in that they are based on discrete point scales where a score 2, for example, does not necessarily indicate twice as much plaque as a score 1 (Quigley and Hein 1962; Silness and L  e 1964; Turesky et al. 1970). The use of means and parametric analyses may be inappropriate, and the clinical relevance of reductions in plaque as a consequence of an intervention may be difficult, if not impossible to determine (McCracken et al. 2006).

Several attempts, using transmission electron microscopy (TEM) (Jentsch et al. 2002) and confocal laser scanning microscopy (CLMS) (Zaura-Arite et al. 2001; Arweiler et al. 2008), have been made to measure the thickness of plaque biofilm. The study by Jentsch et al. (2002) resulted in no significant changes in the biofilm thickness when 72-h-old plaque biofilms were chemically treated with antibacterial mouth rinses according to TEM images that were taken after a very complicated process of sample preparations, including fixation, dehydration, embedding, and sectioning. The details of CLMS images for the calcula-

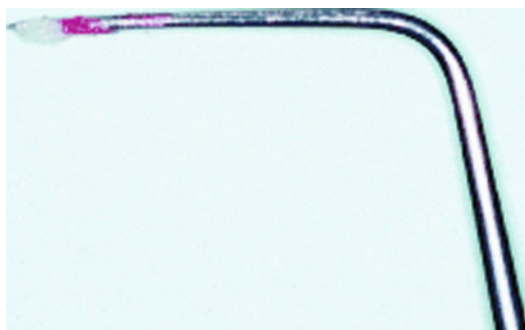


Fig. 18.42 An example of dental plaque removed from a contact point site. The unstained deposit indicates the deposit that would be included as a weight measurement but not included in the index score (McCracken et al. 2006). Reprinted with permission from John Wiley & Sons, Inc.)

tion of biofilm thickness would be also limited in the thinner or the outer region of biofilm, as there is a limit to the scanning depth in the sagittal direction when using CLMS. There have been few trials evaluating the outcomes of mechanical plaque control by directly measuring the thickness of the biofilms that have found a large difference in thickness (Kato et al. 2012).

The antiplaque effect of several therapeutic approaches has been clarified by a reduction in plaque biofilms treated with antimicrobial agents (Helldén et al. 1981), stannous fluoride (Tinanoff et al. 1980), and sugar alcohols (Söderling et al. 1989). The plaque inhibition in these studies was evaluated by measuring the weight of the plaque collected from the affected tooth surfaces (Kato et al. 2012). Studies that evaluated the reduction of plaque biofilm by chemical treatment, using both objective and subjective measures of biofilm (Helldén et al. 1981; Claydon et al. 2004), showed that the objective measurement of plaque weight was superior in performance to the subjective plaque indices (Fig. 18.42). The quantitative or objective evaluation of the amount of plaque was performed using, for example, wet weight, dry weight, the volume of a body, and estimation based on total protein contained (Kato et al. 2012 and references therein). The accuracy of these results, however, would decrease as the thickness of the biofilm grew thinner because accuracy depends on the volume of the samples collected from the tooth

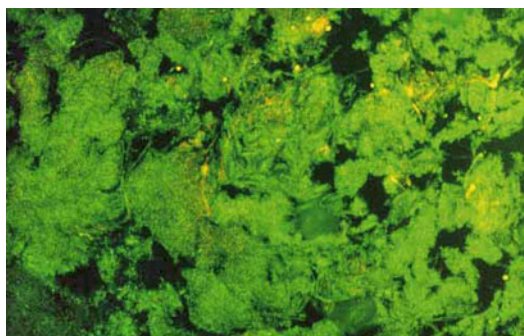


Fig. 18.43 Representative vital fluorescence staining of a plaque sample containing mainly living cells (living cells are stained green) (Sculean et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)

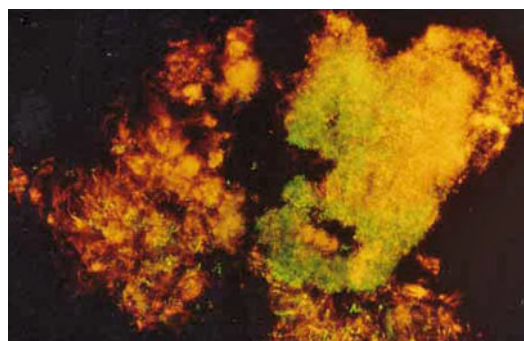


Fig. 18.44 Representative vital fluorescence staining of a plaque sample containing mainly dead cells (dead cells are stained red) (Sculean et al. 2001. Reprinted with permission from John Wiley & Sons, Inc.)

surface. From this point of view, measuring the thickness of plaque biofilm would also be an adequate way to evaluate effectively in a trial the antiplaque effect of using chemical plaque controls (Kato et al. 2012; McCracken et al. 2006).

18.3.19 Bacterial Vitality Within the Dental Plaque

Bacterial vitality within the dental plaque is a qualitative measurement of the antibacterial effect. It is a laboratory parameter and measured by using the vital fluorescence technique (Netuschil et al. 1989; Weiger et al. 1992). Briefly, the technique is based on the use of fluorescein diacetate (FDA) and ethidium bromide (EB). FDA, a fluorescent

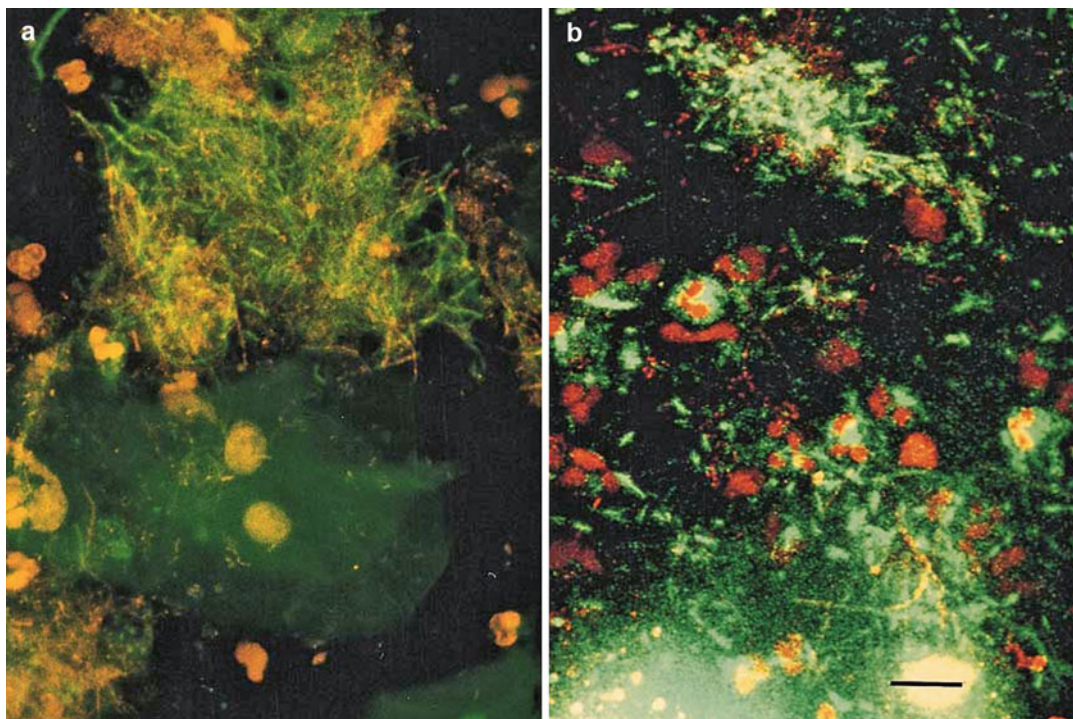


Fig. 18.45 Plaque smears on glass slides after staining with fluorescein diacetate/ethidium bromide. **(a, left)** Conventional epifluorescence microscope (Zeiss photomicroscope III): filamentous green, living plaque microflora together with orange dead bacterial material (upper half) and desquamated cell remnants (faint green bodies) and counterstained, orange cell nuclei (bottom

part). **(b, right)** Confocal laser scanning microscopy: discrimination of vital (green) and dead (red) bacteria and cell nuclei (large red dots). Bar = 10 μ m (objective $\times 63$). Magnification $\times 750$. Different colours in the two illustrations are due to diverse filter characteristics (Netuschil et al. 1998. Reprinted with permission from Elsevier)

dye, is not fluorescent but membrane soluble. In vital cells, it is metabolized to fluorescein which is green and cannot leave the cell. Living cells are stained green (**Fig. 18.43**). Dead cells are not able to metabolize the FDA so that there is no staining. A contra-staining with ethidium bromide (EB) binds to the nucleic acids of dead cells and stains red (**Fig. 18.44**). The method allows living and dead cells to be simultaneously stained (**Fig. 18.45**). Thus, a dichotomous decision living/dead can be made for each single cell. The samples were immediately processed under the fluorescent microscope to enumerate the % of vital flora (VF%) using a counting grid (Brex et al. 1990; Netuschil et al. 1998; Sculean et al. 2001). When the spatial structure of dental biofilms was examined a vital fluorescence technique combined with optical analysis of sections in a confocal laser scanning microscope (CLSM), it was

revealed that in all instances, the bacterial vitality increased from the enamel surface to the central part of the plaque and decreased again in the outer parts of the biofilm (**Fig. 18.46**). The spatial arrangement of the microorganisms in the biofilm showed voids outlined by layers of vital bacteria, which themselves were packed in layers of dead material (**Fig. 18.47**) (Auschill et al. 2001).

This technique was used for the assessment of the vitality of biofilm microorganisms throughout mouthwash application in short-term investigations as well as in experimental gingivitis and long-term studies (Arweiler et al. 2000, 2001a, 2002b, 2003, 2006, 2008; Auschill et al. 2001, 2005; Brex et al. 1990, 1993; Gehlen et al. 2000; König et al. 2002; Lorenz et al. 2009c; McBain et al. 2003; Netuschil et al. 1989, 1995, 1996, 1998; Rundegren et al. 1992; von Ohler et al. 1998; Weiger et al. 1995, 1998).

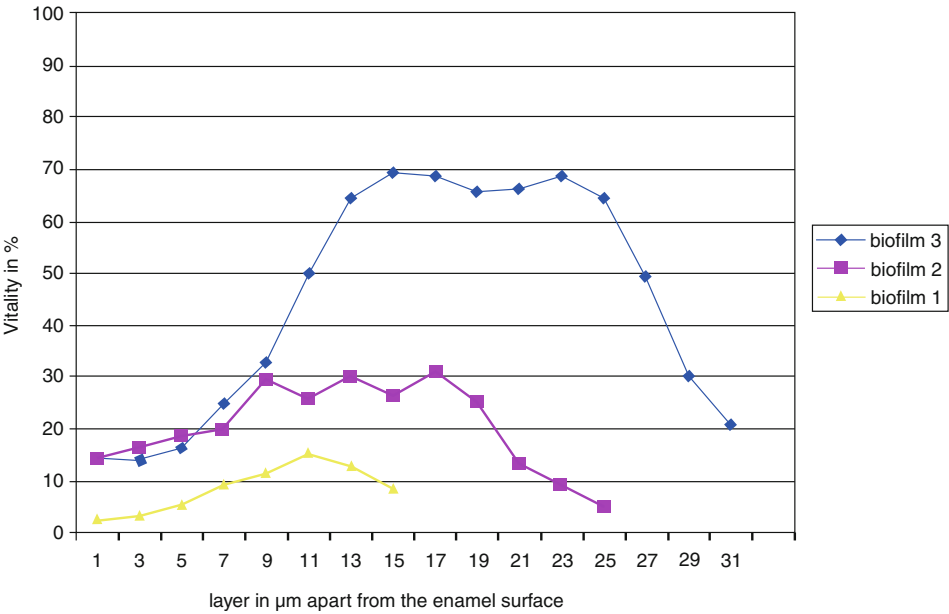


Fig. 18.46 Height and vitality distribution in the different optical layers of three dental biofilms (Auschill et al. 2001. Reprinted with permission from Elsevier)

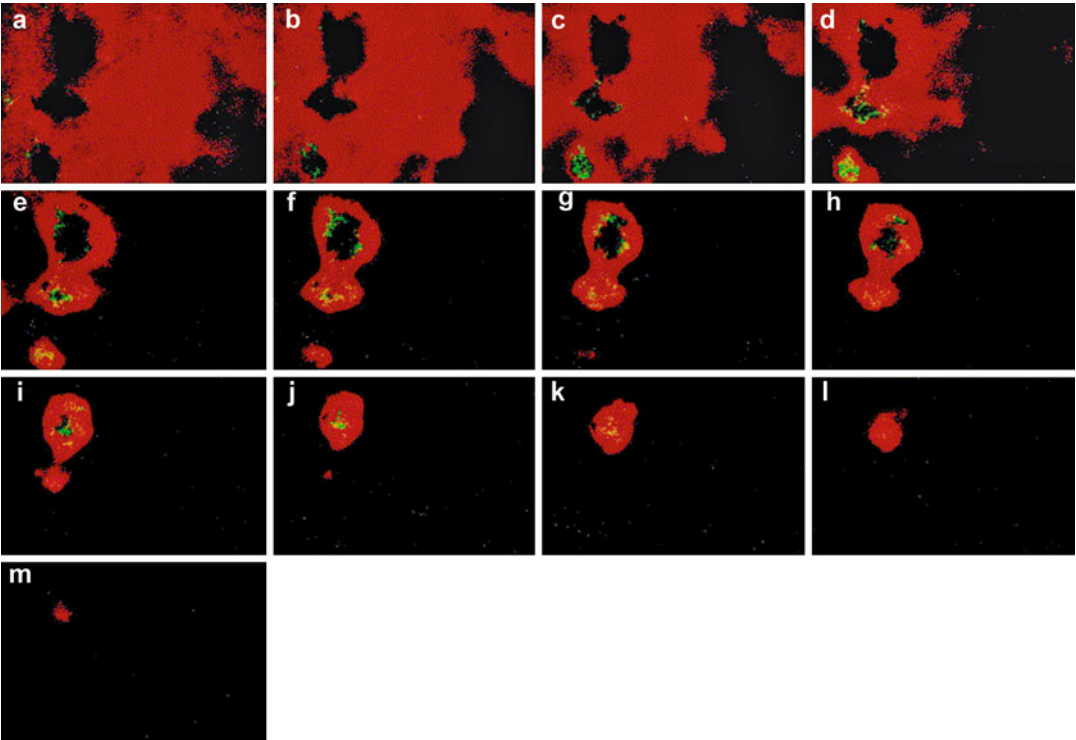


Fig. 18.47 Optical sections (1 μm each) of biofilm 2, vitality fluorescence visualization. First layer starting adjacent to enamel (*bottom, left*), 13th layer 25 μm apart from *bottom (top, left)* (Auschill et al. 2001. Reprinted with permission from Elsevier)

18.3.20 Oral Calculus Index (Greene and Vermillion 1964)

See the oral hygiene index (Greene and Vermillion 1964), composed by the debris index (DI-S) and the calculus index (CI-S).

18.3.21 Calculus Index (Ramfjord 1959)

The following teeth were selected as indicators of the periodontal condition within the dentition: maxillary right first molar 16, maxillary left central incisor 21, maxillary left first bicuspid 24, mandibular left first molar 36, mandibular right central incisor 41, and mandibular right first bicuspid 44. Calculus recording as follows:

Score 0=absence of calculus

Score 1=supragingival calculus extending only slightly below the free gingival margin (not more than 1 mm)

Score 2=moderate amount of supra- and subgingival calculus or subgingival calculus only

Score 3=an abundance of supra and subgingival calculus

The scores on calculus for each individual tooth examined are added and the sum divided by the number of teeth examined to yield the index on calculus (Ramfjord 1959).

18.3.22 Calculus Surface Index (Ennervet et al. 1961)

Ennervet et al. (1961) selected only the four mandibular incisors as the teeth to be graded, based upon the rate of calculus formation and the ease of observing the calculus on these teeth. In applying the scoring method, calculus was considered to be present in any amount, supragingival or subgingival, and it could be detected either visually or by touch. If the examiner was uncertain about the presence of calculus on a given surface, the surface was called calculus free. Each of the mandibular incisors was considered on the basis of 4 surfaces, 2 proximals, 1 labial, and 1 lingual. At examination, a number (zero to four) corresponding to the number of surfaces on which calculus had occurred

was assigned on each tooth. The total number of surfaces on which calculus was detected was considered to be the subject's calculus score and is referred to as the 'calculus surface index' (CSI).

18.3.23 Calculus Rating (Volpe and Manhold 1962)

Volpe and Manhold (1962) presented a method for measuring calculus formation *in vivo* using a coloured periodontal probe. The periodontal probe was placed against the lingual surface of the anterior tooth to be scored with the probe end placed at the most inferior border of any calculus present (supra- or subgingival). With the different colours at the probe end representing units, the amount of calculus present can be measured as:

Score 0=no calculus

Score 1=1 mm of calculus

Score 2=2 mm of calculus

Score 3=3 mm of calculus

The subject's calculus rating is calculated by adding the total number of calculus units measured on the lingual surfaces of the lower anterior teeth (Volpe and Manhold 1962).

18.4 Miscellaneous Indices

18.4.1 Retention Index System (Löe 1967)

The purpose of creating a retention index system was to introduce a system for the assessment of the main retentive factors and which expressed the quality of the tooth surface (degree of roughness) adjacent to the gingival tissues (Löe 1967). The scores are as follows:

Score 0=no caries, no calculus, no imperfect margin of dental; restoration in a gingival location

Score 1=supragingival cavity, calculus, or imperfect margin of dental restoration

Score 2=subgingival cavity, calculus, or imperfect margin of dental restoration

Score 3=large cavity, abundance of calculus, or glossy insufficient marginal fit of dental restoration in a supra- and/or subgingival location (Löe 1967)

18.4.2 Mobility Index (Ramfjord 1959)

The following teeth were selected as indicators of the periodontal condition within the dentition: maxillary right first molar 16, maxillary left central incisor 21, maxillary left first bicuspid 24, mandibular left first molar 36, mandibular right central incisor 41, and mandibular right first bicuspid 44. Record mobility as follows:

Score M 0=physiologic mobility; firm tooth

Score M 1=slightly increased mobility

Score M 2=definite to considerable increase in mobility but no impairment of function

Score M 3=extreme mobility; a 'loose tooth' that cannot be used for normal function

The scores for mobility for each individual tooth examined are added and the sum divided by the number of teeth examined to yield the index of mobility.

categorized by 1 of 32 general dental practitioners using the PIT technique upon the six test teeth and subsequently viewing the bitewing radiographs. The results showed that, when used without bitewings as an adjunct, the PIT technique showed a specificity of 95–100% but a low sensitivity in three groups (36%, 52%, and 59%) and of 94% in the fourth group who had in fact been referred for periodontal treatment. When bitewing radiographs were used as an adjunct to periodontal probing, the PIT technique showed a specificity of 100% in each group and a sensitivity of 89–100% when screening for pockets and/or bone loss. It was concluded that, in the subjects studied, the PIT technique when used in conjunction with bitewing radiography provided simple, rapid, and reliable periodontal screening (Eaton and Woodman 1989).

18.5 Treatment Needs Indices

18.5.1 Periodontal Index for Treatment (Eaton and Woodman 1989)

Eaton and Woodman (1981–1985) developed the periodontal index for treatment (PIT) in which the clinical assessment of six teeth (all first molars and maxillary right and mandibular left central incisors) is completed with a specially designed periodontal probe (PIT probe with markings at 4, 6, 8, and 11 mm and a 0.5-mm ball tip). The maxillary sites are probed on the palatal side while the mandibular sites are probed on the buccal side, and scores are given from 0 to 3. The overall patient score (PIT score) is recorded as the highest score of the six test teeth (Dhingra and Vandana 2011).

Eaton and Woodman (1989) examined 406 UK Royal Navy and Royal Air force personnel in 1989 using this index with and without bitewing radiographs. The subjects were categorized as periodontally healthy (PIT 0), with gingivitis but no pocketing exceeding 4 mm (PIT 1), with pocketing 4–5 mm (PIT 2), or with 6 mm+ pocketing (PIT 3). The subjects were then examined and

18.5.2 Community Periodontal Index of Treatment Needs (CPITN)

The recommendation for scoring the community periodontal index of treatment needs (CPITN) was agreed at the meeting of the joint FDI/WHO Working Group in 1981. The index was primarily designed to assess periodontal treatment needs rather than periodontal status, that is, recession of the gingival margin and alveolar bone (Ainamo et al. 1982).

The periodontal treatment needs are recorded for sextants, that is, sixths of the dentition. Third molars are not included, except where they are functioning in the place of second molars. The treatment need in a sextant is recorded only when two or more teeth are present and not indicated for extraction. If only one functioning tooth remains in a sextant, it is included in the adjoining sextant. Missing sextants are indicated with a diagonal line through the appropriate box (Ainamo et al. 1982).

In epidemiological surveys, assessing the periodontal treatment needs of a population the recordings per sextant are based on findings from specified index teeth. The index teeth to be examined are 17, 16, 11, 26, 27, 47, 46, 31, 36, and 37. Although 10 index teeth are examined, only 6

recordings, one relating to each sextant, are made. When both or one of the designated molar teeth are present, the worst finding from these tooth surfaces is recorded for the sextant. If no index teeth are present in a sextant qualifying for examination, all the remaining teeth in that sextant are examined (Ainamo et al. 1982).

To apply the trial method of measurement, a special probe was developed. The WHO periodontal probe has a 'ball-tip' end for easier detection of subgingival calculus and to avoid false reading for over-measurement. The colour-coded area from 3.5 to 5.5 mm greatly facilitates rapid reading of pocket depth.

In assessing treatment needs, the presence of the following is determined for each sextant in the sequence indicated:

Score 0=no sign of disease

Score 1=gingival bleeding after gentle probing

Score 2=supra- or subgingival calculus

Score 3=pathologic pockets 4 or 5 mm deep

Score 4=pathologic pockets 6 mm or deeper

Subjects are classified into the different treatment need categories according to the highest score recorded during the examination:

Code 0: no treatment

Code 1: improvement in oral hygiene

Codes 2 and 3: improvement in oral hygiene+scaling

Code 4: improvement in oral hygiene+scaling+complex treatment, which may involve deep scaling and root planning under local anaesthesia, or require surgical exposure of the infected root surface in order to gain the access needed to clean it.

The CPITN was designed for rapid and practical assessment of various periodontal treatment needs in population surveys and for initial screening of patients attending for regular dental care. It must be emphasized that the CPITN does not provide an assessment of past periodontal experience. It does not record the position of the gingival margin, that is, the degree of recession, and, hence, except where pocketing is present, the level of the alveolar bone (Ainamo et al. 1982). Recently, the two versions of CPITN for periodontitis diagnosis were validated in a sample of 400 individuals underwent full-mouth periodontal examination including clinical attachment

loss, periodontal probing depth, and subgingival calculus (gold standard). The results showed 58% sensitivity for full CPITN and 80.6% specificity. Positive and negative predictive values were 87% and 46.3%, respectively. According to the test, estimated periodontitis prevalence was 46%, while the figure obtained with the gold standard was 69%. The partial version of the CPITN showed 50% sensitivity and 87.1% specificity. Positive and negative predictive values were 89.6% and 43.9%, respectively. Estimated periodontitis prevalence, through partial CPITN, was 30.5%. Adjusted global agreement (kappa) for partial and full CPITN was 0.32 and 0.29, respectively. As a conclusion, both total and partial CPITN failed to reflect the real periodontal status of the sample (Bassani et al. 2006). Even if the CPITN index was extensively used in epidemiological studies, this index has a very limited use also for expressing the severity and prevalence of periodontal destruction in a population because it is not site-based and it severely under- and overestimates the destructive aspects of periodontitis (like the prevalence of deep pockets, calculus, bleeding) (Baelum and Papapanou 1996; Beck and Loe 1993). As Leroy et al. (2010) summarized, the CPITN index is based on a hierarchical concept of the progression of periodontitis which implies that a tooth with a score of 3 or 4 (a pocket present) should also have calculus present (score 2) and bleeding (score 1). The validity of this assumption has been challenged (Benigeri et al. 2000; Lewis et al. 1994). In a Norwegian population, 30% of teeth with calculus did not present with bleeding and 25% with deep pockets (score 4), and bleeding did not have any calculus present (Grytten et al. 1989). In a Japanese population, bleeding was absent in 47.5% of sextants with a CPITN score of 2 for calculus (Takahashi et al. 1988). In Hong Kong, Holmgren and Corbet (1990) reported a similar finding and concluded that the presence of calculus without bleeding at a sizeable proportion of index teeth without pockets questioned the assumption of the CPITN that there is a close concordance between calculus and periodontal inflammation (Holmgren and Corbet 1990). Further limitations of CPITN are that it does not measure tooth mobility or attach-

ment loss (Cutress et al. 1987; Baelum et al. 1988) or furcation involvement. Gera (2000) has suggested that in populations with access to periodontal care, teeth may experience considerable gingival recession following therapy and have minimal pocket depth leading to CPITN scores of 0 or 1, when there has actually been considerable past attachment loss, thus leading to underestimates of the extent and severity of periodontal destruction that has previously occurred in their mouths (Gera 2000). There are also doubts about the ability of any technique that examines the periodontium around just a few teeth to reflect the true state of periodontal health or disease in the mouth concerned. These limitations have led to the value of the CPITN being questioned as a reliable epidemiological tool (Baelum et al. 1995; Leroy et al. 2010).

18.5.3 The Community Periodontal Index (CPI) INDEX (World Health Organization 1997)

Later, CPITN index (Ainamo et al. 1982) was modified to CPI (World Health Organization 1997). Nevertheless, the instrument's diagnostic criteria were basically unaffected by the modifications (Bassani et al. 2006). The major advantages of the CPI are simplicity, speed, reproducibility, and international uniformity (Petersen and Ogawa 2005). The limitations of CPI are its inability to provide an adequate assessment of prevalence of periodontal disease (Vettore et al. 2007).

The community periodontal index has been criticized for being an out-of-date paradigm to assess disease; does not include the cumulative manifestations of periodontal destruction such as clinical attachment loss, gingival recession, and marginal bone loss; and, because of gingival recession, increasingly underestimates disease severity with increasing age. In adolescent populations, the partial recording system tends to distort estimates of prevalence and severity (Jenkins and Papapanou 2001). Furthermore, in young people with low prevalence of periodontal pockets, the validity of the hierarchical design has been questioned and would not lead to the underestimation

of dental calculus prevalence. However, the same is not true for bleeding, as the dental calculus record (CPI=2) does not enable the identification of the concomitant bleeding (CPI=1) in the same sextant (Antunes et al. 2008). This observation is consistent with the reporting of higher prevalence of gingival bleeding in two other surveys among Brazilian adolescents, which modified the diagnostic criteria established by the WHO (Antunes et al. 2008; Gesser et al. 2001).

18.5.4 Periodontal Screening and Recording (PSR) Index (Landry and Jean 2002)

The next periodontal index to be developed was the periodontal screening and recording index (PSR). The concept of periodontal screening first arose in 1988 when the American Academy of Periodontology (AAP) approved the development of a Periodontal Disease Detection Day, in collaboration with the American Dental Association (ADA). The AAP committee decided to use a modified CPITN procedure for periodontal screening, and a position statement on the use of this screening procedure was developed which later evolved into the professional education brochure on a Periodontal Screening and Recording system. Subsequently, in 1990, the PSR was reviewed and further modified by ADA committee, and, in 1991, it was endorsed by the ADA as a useful tool for assessing the periodontal status of patients. In 1992, Procter and Gamble Company (P&G) became the official sponsor of PSR system and helped in active distribution of PSR training programme kits and promotion of the PSR system among the dentists in USA from 1992 to 1994. Subsequently, the PSR system became highly popular in the USA, and the Canadian Dental Association and Canadian Periodontist Association also adopted the index in August, 1995 (Dhingra and Vandana 2011; Nasi, 1994; Landry and Jean 2002).

Apart from one major difference, the PSR index is virtually identical to CPITN. The two indices use a common evaluation method based on the following three periodontal disease indica-

tors: gingival bleeding on probing, calculus accumulation, and probing depth. In addition, the PSR index provides a more detailed picture of periodontal status by recording the presence of furcation involvement, tooth mobility, mucogingival problems, and gingival recessions exceeding 3.5 mm. When at least one of the above conditions is present, an asterisk (*) is recorded with the PSR score for that given sextant (Landry and Jean 2002).

For both the PSR and CPITN indices, parameters are evaluated using a periodontal probe similar to the one introduced by WHO in 1982. The instrument must have a 0.5-mm-diameter ball tip and a coloured band extending from 3.5 to 5.5 mm from the tip. The ball tip distributes the force applied during probing over a larger surface area, thus reducing the risk of injury at the epithelial junction and making the procedure more comfortable for patient. Additionally, the ball ended tip is designed to move the shank of the probe away from the tooth in order to make identification of subgingival calculus and overhanging restorations easier. The coloured band allows a quick estimate to be made of pocket depth and the degree of gingival recession, eliminating the need for recording precise measurements. The probe is thus designed to rapidly differentiate 'normal' from 'abnormal' status to provide an estimate of an individual's periodontal treatment needs (Landry and Jean 2002).

Information is collected and scored by dividing the mouth in to six sextants. The first sextant includes teeth 18–14 according to the WHO classification; the second, teeth 13–23; and so forth. Each tooth in a sextant is probed at six different sites: mesio-buccal, mid-buccal, disto-buccal, and the corresponding lingual and palatal sites. To detect pathological conditions without producing pain or false results, a maximum probing force of 20 g is recommended. Each tooth is scored from code 0 to code 4, but only the highest score of the sextant is recorded. In addition, sextants with fewer than two teeth are scored with an 'X' and are not considered in the overall evaluation. However, if a sextant has only one functional tooth, it is included in the preceding sextant. The periodontal health of a sextant is thus

assessed by the highest score, which corresponds to the most important clinical sign and designates a recommended treatment plan (Landry and Jean 2002).

The American Dental Association and the American Academy of Periodontology recommend that PSR be conducted by dentists for all patients as an integral part of oral examinations. The benefits of the PSR index are (<http://www.armydentalscaresystem.army.mil/usarc/training/PSRappF.pdf>):

1. Early detection: PSR includes evaluation of all sites at periodontal risk. For this reason, it is a highly sensitive technique for detecting deviations from periodontal health and a uniquely appropriate screening tool for periodontal diseases that are, by nature, site specific and episodic.
2. Speed: Once learned, PSR takes only a few minutes to conduct for each patient. It can be readily incorporated into routine oral examinations without lengthening appointment time.
3. Simplicity: PSR is easy to administer and comprehend. The simplicity of the scoring system aids in monitoring patients and helps patients understand their periodontal status.
4. Cost-effectiveness: PSR utilizes a simple periodontal probe designed specifically for use with this screening system. It does not require the use of expensive equipment.
5. Recording ease: Documentation for PSR requires the recording of six numerical scores, one for each sextant of the mouth. It does not require extensive charting or lengthy narrative explanation.
6. Risk management: Proper, consistent, and documented use of PSR shows that the dentist is evaluating a patient's periodontal status and satisfies dental-legal requirements in the area of monitoring and record keeping.

The limitations of the PSR index are as follows (<http://www.ada.org/3070.aspx#benefits>):

1. It is a screening, not a substitute for a comprehensive periodontal examination.
2. Patients in the maintenance phase of periodontal treatment require periodic comprehensive periodontal examinations.
3. Intended primarily for adults.

References

- Abrahamsson I, Soldini C. Probe penetration in periodontal and peri-implant tissues. An experimental study in the beagle dog. *Clin Oral Implants Res*. 2006;17:601–5.
- Adams RA, Nystrom GP. A periodontitis severity index. *J Periodontol*. 1986;57:176–9.
- Addy M, Willis L, Moran J. Effect of toothpaste rinses compared with chlorhexidine on plaque formation during a 4-day period. *J Clin Periodontol*. 1983;10:89–99.
- Addy M, Renton-Harper P, Newcombe R. Plaque regrowth studies: discriminatory power of plaque index compared to plaque area. *J Clin Periodontol*. 1999;26:110–2.
- Ainamo J, Bay I. Problems and proposals for recording gingivitis and plaque. *Int Dent J*. 1975;25:229–35.
- Ainamo J, Barnes D, Beagrie G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) community periodontal index of treatment needs (CPITN). *Int Dent J*. 1982;32:281–91.
- Akhter R, Hannan MA, Okhubo R, Morita M. Relationship between stress factor and periodontal disease in a rural area population in Japan. *Eur J Med Res*. 2005;10:352–7.
- Albandar JM. Periodontal diseases in North America. *Periodontol 2000*. 2002;29:31–69.
- Albandar JM. Underestimation of periodontitis in NHANES surveys. *J Periodontol*. 2011;82:337–41.
- Albandar JM, Brunelle JA, Kingman A. Destructive periodontal disease in adults 30 years of age and older in the United States, 1988–1994. *J Periodontol*. 1999;70:13–29.
- Anonymous. Parameter on chronic periodontitis with slight to moderate loss of periodontal support. American Academy of Periodontology. *J Periodontol*. 2000;71(5 Suppl):853–5.
- Antunes JL, Peres MA, Frias AC, Crosato EM, Biazevic MG. Gingival health of adolescents and the utilization of dental services, state of São Paulo, Brazil. *Rev Saude Publica*. 2008;42:191–9.
- Armitage GC, Jeffcoat MK, Chadwick DE, Taggart Jr EJ, Numabe Y, Landis JR, Weaver SL, Sharp TJ. Longitudinal evaluation of elastase as a marker for the progression of periodontitis. *J Periodontol*. 1994;65:120–8.
- Arweiler NB, Donos N, Netuschil L, Reich E, Sculean A. Clinical and antibacterial effect of tea tree oil – a pilot study. *Clin Oral Investig*. 2000;4:70–3.
- Arweiler NB, Netuschil L, Reich E. Alcohol-free mouthrinse solutions to reduce supragingival plaque regrowth and vitality. A controlled clinical study. *J Clin Periodontol*. 2001a;28:168–74.
- Arweiler NB, Reich E, Brex M, Dörner M, Netuschil L. Influence of ethanol on the antibacterial and anti-plaque efficacy of an amine fluoride/stannous fluoride mouthrinse solution. *J de Parodontologie de l'Implantologie orale*. 2001b;20:331–40.
- Arweiler NB, Auschill TM, Reich E, Netuschil L. Substantivity of toothpaste slurries and their effect on reestablishment of the dental biofilm. *J Clin Periodontol*. 2002a;29:615–21.
- Arweiler NB, Henning G, Reich E, Netuschil L. Effect of an amine-fluoride-triclosan mouthrinse on plaque regrowth and biofilm vitality. *J Clin Periodontol*. 2002b;29:358–63.
- Arweiler NB, Auschill TM, Baguley N, Netuschil L, Sculean A. Efficacy of an amine fluoride-triclosan mouthrinse as compared to the individual active ingredients. *J Clin Periodontol*. 2003;30:192–6.
- Arweiler NB, Boehnke N, Sculean A, Hellwig E, Auschill TM. Differences in efficacy of two commercial 0.2% chlorhexidine mouthrinse solutions: a 4-day plaque regrowth study. *J Clin Periodontol*. 2006;33:334–9.
- Arweiler NB, Lenz R, Sculean A, Al-Ahmad A, Hellwig E, Auschill TM. Effect of food preservatives on in situ biofilm formation. *Clin Oral Investig*. 2008;12:203–8.
- Auschill TM, Arweiler NB, Netuschil L, Brex M, Reich E, Sculean A. Spatial distribution of vital and dead microorganisms in dental biofilms. *Arch Oral Biol*. 2001;46:471–6.
- Auschill TM, Hein N, Hellwig E, Follo M, Sculean A, Arweiler NB. Effect of two antimicrobial agents on early in situ biofilm formation. *J Clin Periodontol*. 2005;32:147–52.
- Baelum V, Papapanou PN. CPITN and the epidemiology of periodontal disease. *Community Dent Oral Epidemiol*. 1996;24:367–8.
- Baelum V, Fejerskov O, Manji F. Periodontal diseases in adult Kenyans. *J Clin Periodontol*. 1988;15:445–52.
- Baelum V, Manji F, Wanzala P, Fejerskov O. Relationship between CPITN and periodontal attachment loss findings in an adult population. *J Clin Periodontol*. 1995;22:146–52.
- Barendregt DS, Timmerman MF, Van der Velden U, Van der Weijden GA. Comparison of the bleeding on marginal probing index and the Eastman interdental bleeding index as indicators of gingivitis. *J Clin Periodontol*. 2002;29:195–200.
- Barnett ML. Suitability of gingival indices for use in therapeutic trials. Is bleeding a sine qua non? *J Clin Periodontol*. 1996;23:582–6.
- Barnett ML, Ciano SG, Mather ML. The modified papillary bleeding index: comparison with gingival index during the resolution of gingivitis. *J Prev Dent*. 1980;6:135–8.
- Bassani DG, da Silva CM, Oppermann RV. Validity of the “Community Periodontal Index of Treatment Needs” (CPITN) for population periodontitis screening. *Cad Saude Publica*. 2006;22:277–83.
- Beck JD, Loe H. Epidemiological principles in studying periodontal diseases. *Periodontol 2000*. 1993;2:34–45.
- Benamghar L, Penaud J, Kaminsky P, Abt F, Martin J. Comparison of gingival index and sulcus bleeding index as indicators of periodontal status. *Bull World Health Organ*. 1982;60:147–51.
- Benigeri M, Brodeur JM, Payette M, Charbonneau A, Ismail AI. Community periodontal index of treatment needs and prevalence of periodontal conditions. *J Clin Periodontol*. 2000;27:308–12.
- Benson BJ, Henyon G, Grossman E, Mankodi S, Sharma NC. Development and verification of the proximal/marginal plaque index. *J Clin Dent*. 1993;4:14–20.

- Bessa Rebelo MA, Adriana Correia de Queiroz A. Gingival indices: state of art. In: Panagakos FS, Davies RM, editors. *Gingival diseases – their aetiology, prevention and treatment*. InTech; 2011. p. 41–54. <http://www.intechopen.com/profiles/65936/maria-augusta-rebelo>
- Birkeland JM, Jorkjend L. The influence of examination on the assessment of the intra-examiner error by using the plaque and gingival index systems. *Community Dent Oral Epidemiol*. 1975;3:214–6.
- Bliden TM, Caton JG, Proskin HM, Stein SH, Wagener CJ. Examiner reliability for an invasive gingival bleeding index. *J Clin Periodontol*. 1992;19:262–7.
- Borges-Yáñez SA, Maupomé G, Jiménez-García G. Validity and reliability of partial examination to assess severe periodontitis. *J Clin Periodontol*. 2004;31:112–8.
- Borrell LN, Beck JD, Heiss G. Socioeconomic disadvantage and periodontal disease: the Dental Atherosclerosis Risk in Communities study. *Am J Public Health*. 2006;96:332–9.
- Brexc M, Netuschil L, Reichert B, Schreil G. Efficacy of Listerine, Meridol and chlorhexidine mouthrinses on plaque, gingivitis and plaque bacteria vitality. *J Clin Periodontol*. 1990;17:292–7.
- Brexc M, MacDonald LL, Legary K, Cheang M, Forgay MGE. Long-term effects of Meridol and chlorhexidine mouthrinses on plaque, gingivitis, staining, and bacterial vitality. *J Dent Res*. 1993;72:1194–7.
- Brothwell D, Ghiabi E. Periodontal health status of the Sandy Bay First Nation in Manitoba, Canada. *Int J Circumpolar Health*. 2009;68:23–33.
- Butler A, Targett D, Bosma ML. Maintenance of gingival health – a measure based on clinical indices. *Int Dent J*. 2011;61 Suppl 3:28–32.
- Camgoz M, Gurgan CA, Kajiwaru K, Kawamura M. Dental students' ability to assess gingival health status with DAAGS software. *J Dent Educ*. 2008;72:59–66.
- Camgoz M, Gurgan CA, Akkaya M. Turkish dental students' and dentists' ability to assess gingival health status with DAAGS software. *J Dent Educ*. 2011;75:1127–32.
- Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. *J Clin Periodontol*. 1986;13:500–5.
- Carter HG, Barnes GP. The gingival bleeding index. *J Periodontol*. 1974;45:801–5.
- Carter K, Landini G, Walmsley AD. Automated quantification of dental plaque accumulation using digital imaging. *J Dent*. 2004;32:623–8.
- Caton JG, Polson AM. The interdental bleeding index: a simplified procedure for monitoring gingival health. *Compend Contin Educ Dent*. 1985;6(88):90–2.
- Caton J, Bouwsma O, Polson A, Espeland M. Effects of personal oral hygiene and subgingival scaling on bleeding interdental gingiva. *J Periodontol*. 1989;60:84–90.
- Chaves ES, Wood RC, Jones AA, Newbold DA, Manwell MA, Kornman KS. Relationship of “bleeding on probing” and “gingival index bleeding” as clinical parameters of gingival inflammation. *J Clin Periodontol*. 1993;20:139–43.
- Chawla TN. Prevalence of periodontal disease in urban Lucknow (India) using Ramfjord's Technique. *J All India Dent Assoc*. 1963;35:151–9.
- Ciancio SG. Current status of indices of gingivitis. *J Clin Periodontol*. 1986;13(375–8):381–2.
- Claydon N, Yates R, Labello R, Hall C, Armstrong R, Addy M. A methodology using subjective and objective measures to compare plaque inhibition by toothpastes. *J Clin Periodontol*. 2004;31:1106–9.
- Clements KM, Bollen AM, Huang G, King G, Hujoel P, Ma T. Activation time and material stiffness of sequential removable orthodontic appliances. Part 2: dental improvements. *Am J Orthod Dentofacial Orthop*. 2003;124:502–8.
- Cugini M, Thompson M, Warren PR. Correlations between two plaque indices in assessment of toothbrush effectiveness. *J Contemp Dent Pract*. 2006;7(5):1–9.
- Cutress TW, Ainamo J, Sardo-Infirri J. The community periodontal index of treatment needs (CPITN) procedure for population groups and individuals. *Int Dent J*. 1987;37:222–33.
- Dababneh RH, Khouri AT, Smith RG, Addy M. A new method of plaque scoring: a laboratory comparison with other plaque indices. *J Clin Periodontol*. 2002a;29:832–7.
- Dababneh RH, Khouri AT, Smith RG, Addy M. Correlation and examiner agreement between a new method of plaque scoring and a popular established plaque index, modelled in vitro. *J Clin Periodontol*. 2002b;29:1107–11.
- Despeignes JR. Variation in the area of intraperiodontal surfaces of human tooth roots, in relation to their depth. Thesis Paris 1970. *J Periodontol*. 1979;50:630–5.
- Dhingra K, Vandana KL. Indices for measuring periodontitis: a literature review. *Int Dent J*. 2011;61:76–84.
- Dunning JM, Leach LB. Gingival-bone count: a method for epidemiological study of periodontal disease. *J Dent Res*. 1960;39:506–13.
- Dye BA, Barker LK, Selwitz RH, Lewis BG, Wu T, Fryar CD, Ostchega Y, Beltran ED, Ley E. Overview and quality assurance for the National Health and Nutrition Examination Survey (NHANES) oral health component, 1999–2002. *Community Dent Oral Epidemiol*. 2007;35:140–51.
- Dye BA, Thornton-Evans G. A brief history of national surveillance efforts for periodontal disease in the United States. *J Periodontol*. 2007;78:1373–9.
- Dye BA, Hirsch R, Brody DJ. The relationship between blood lead levels and periodontal bone loss in the United States, 1988–1994. *Environ Health Perspect*. 2002a;110:997–1002.
- Dye BA, Kruszon-Moran D, McQuillan G. The relationship between periodontal disease attributes and *Helicobacter pylori* infection among adults in the United States. *Am J Public Health*. 2002b;92:1809–15.
- Eaton KA, Woodman AJ. Evaluation of simple periodontal screening technique currently used in UK armed forces. *Community Dent Oral Epidemiol*. 1989;17:190–5.
- Egelberg J. *Oral hygiene methods the scientific way (synopses of clinical studies)* 1st ed. Malmö: Odonto Science Sweden; 1999. p. 271–8.

- Eickholz P, Hausmann E. Accuracy of radiographic assessment of interproximal bone loss in intrabony defects using linear measurements. *Eur J Oral Sci.* 2000;108:70–3.
- Ekanayaka ANI, Sheiham A. Assessing the periodontal treatment needs of a population. *J Clin Periodontol.* 1979;6:150–7.
- Elliott JR, Bowers GM, Clemmer BA, Rovelstad GH. Evaluation of an oral physiotherapy center in the reduction of bacterial plaque and periodontal disease. *J Periodontol.* 1972;43:221–4.
- Ellis JS, Seymour RA, Robertson P, Butler TJ, Thomason JM. Photographic scoring of gingival overgrowth. *J Clin Periodontol.* 2001;28:81–5.
- Engelberger T, Hefti A, Kallenberger A, Rateitschak KH. Correlations among papilla bleeding index, other clinical indices and histologically determined inflammation of gingival papilla. *J Clin Periodontol.* 1983;10:579–89.
- Ennervet J, Sturzenberger OP, Radike AW. The calculus surface index method for scoring clinical calculus studies. *J Periodontol.* 1961;32:54–7.
- Fischman SL. Clinical index systems used to assess the efficacy of mouthrinses on plaque and gingivitis. *J Clin Periodontol.* 1988;15:506–10.
- Fischman S, Cancro LP, Pretara-Spanedda P, Jacobs D. Distal mesial plaque index. A technique for assessing dental plaque about the gingiva. *Dent Hyg (Chic).* 1987;61:404–9.
- Galgut P. A comparison of different indices used in the clinical assessment of plaque and gingival bleeding. *Clin Oral Investig.* 1999;3:96–9.
- Garg S, Kapoor KK. The quantitative gingival bleeding index. *J Indian Dent Assoc.* 1985;57:112–3.
- Gehlen I, Netuschil L, Georg T, Reich E, Berg R, Katsaros C. The influence of a 0.2% chlorhexidine mouthrinse on plaque regrowth in orthodontic patients. *J Orofac Orthop.* 2000;61:138–48.
- Gera I. Periodontal treatment needs in Central and Eastern Europe. *J Int Acad Periodontol.* 2000;2:120–8.
- Gesser HC, Peres MA, Marcenés W. Condições gengivais e periodontais associadas a fatores socioeconômicos. *Rev Saude Publica.* 2001;35:289–93.
- Greene JC, Vermillion JR. The simplified oral hygiene index. *J Am Dent Assoc.* 1964;68:7–29.
- Grossman FD, Fedi PF. Navy periodontal screening examination. *J Am Soc Prev Dent.* 1974;3: 41–45.
- Grytten J, Holst D, Gjermo P. Validity of CPITN's hierarchical scoring method for describing the prevalence of periodontal conditions. *Community Dent Oral Epidemiol.* 1989;17:300–3.
- Hancock EB, Wirthlin Jr MR. An evaluation of the navy periodontal screening examination. *J Periodontol.* 1977;48:63–6.
- Hazen SP. Indices for the measurement of gingival inflammation in clinical studies of oral hygiene and periodontal disease. *J Periodontol Res Suppl.* 1974;14:61–77.
- Helldén L, Camosci D, Hock J, Tinanoff N. Clinical study to compare the effect of stannous fluoride and chlorhexidine mouthrinses on plaque formation. *J Clin Periodontol.* 1981;8:12–6.
- Hofer D, Sahrman P, Attin T, Schmidlin PR. Comparison of marginal bleeding using a periodontal probe or an interdental brush as indicators of gingivitis. *Int J Dent Hyg.* 2011;9:211–5.
- Holmgren CJ, Corbet EF. Relationship between periodontal parameters and CPITN scores. *Community Dent Oral Epidemiol.* 1990;18:322–3.
- Hujoel PP. A meta-analysis of normal ranges for root surface areas of the permanent dentition. *J Clin Periodontol.* 1994;21:225–9.
- Hujoel PP, White BA, García RI, Listgarten MA. The dentogingival epithelial surface area revisited. *J Periodontol Res.* 2001;36:48–55.
- Hunt RJ, Fann SJ. Effect of examining half the teeth in a partial periodontal recording of older adults. *J Dent Res.* 1991;70:1380–5.
- Janssen PTM, Faber JAJ, van Palenstein Helderman WH. Reproducibility of bleeding tendency measurements and the reproducibility of mouth bleeding scores for the individual patient. *J Periodontol Res.* 1986;21: 653–9.
- Jenkins WMM, Papapanou PN. Epidemiology of periodontal disease in children and adolescents. *Periodontol* 2000. 2001;26:16–32.
- Jentsch H, Hombach A, Beetke E, Jonas L. Quantitative transmission electron microscopic study of dental plaque – an in vivo study with different mouthrinses. *Ultrastruct Pathol.* 2002;26:309–13.
- Jovanovski V, Lynch E. Analysis of the morphology of oral structures from 3-D co-ordinate data. *Monogr Oral Sci.* 2000;17:73–129.
- Kato K, Tamura K, Nakagaki H. Quantitative evaluation of the oral biofilm-removing capacity of a dental water jet using an electron-probe microanalyzer. *Arch Oral Biol.* 2012;57(1):30–5.
- Kawamura M, Fukuda S, Inoue C, Sasahara H, Iwamoto Y. The validity and reproducibility of an oral rating index as a measurement of gingival health care and oral hygiene level in adults. *J Clin Periodontol.* 2000;27: 411–6.
- Klages U, Weber AG, Wehrbein H. Approximal plaque and gingival sulcus bleeding in routine dental care patients: relations to life stress, somatization and depression. *J Clin Periodontol.* 2005;32:575–82.
- Klukowska M, Bader A, Erbe C, Bellamy P, White DJ, Anastasia MK, Wehrbein H. Plaque levels of patients with fixed orthodontic appliances measured by digital plaque image analysis. *Am J Orthod Dentofacial Orthop.* 2011;139:e463–70.
- König J, Storcks V, Kocher T, Bössmann K, Plagmann H-C. Antiplatelet effect of tempered 0.2% chlorhexidine rinse: an in vivo study. *J Clin Periodontol.* 2002; 29:207–10.
- Kraus BS, Jordan RE, Abrams L. Dental anatomy and occlusion. A study of the masticatory system. Baltimore/London: Williams & Wilkins; 1969.
- Landry RG, Jean M. Periodontal screening and recording (PSR) index: precursors, utility and limitations in a clinical setting. *Int Dent J.* 2002;52: 35–40.

- Lang NP. Commonly used indices to assess oral hygiene and gingival and periodontal health and disease. In: Lang NP, editor. *Proceedings of the European workshop on mechanical plaque control – status of the art and science of dental plaque control: castle of Münchenwiler, Berne, 9–12 May 1998*. Chicago/Berlin/London/Tokyo/Paris/Barcelona: Quintessence Publ.
- Lange DE. Accessory treatment in systemic periodontitis management. *Zahnärztl Welt*. 1975;8:366–74.
- Lavigne SE. Periodontal assessment in Hodges KO.: concepts in nonsurgical periodontal therapy. Albany/Bonn/Boston: Delmar Publishers; 1998. p. 53–87.
- Leroy R, Eaton KA, Savage A. Methodological issues in epidemiological studies of periodontitis—how can it be improved? *BMC Oral Health*. 2010;10:8.
- Lewis JM, Morgan MV, Wright FA. The validity of the CPITN scoring and presentation method for measuring periodontal conditions. *J Clin Periodontol*. 1994;21:1–6.
- Lobene RR, Weatherford T, Ross NM, Lamm RA, Menaker L. A modified gingival index for use in clinical trials. *Clin Prev Dent*. 1986;8:3–6.
- Löe H. The gingival index, the plaque index and the retention index systems. *J Periodontol*. 1967;38:38–44.
- Loesche WJ. Clinical and microbiological aspects of chemotherapeutic agents used according to the specific plaque hypothesis. *J Dent Res*. 1979;58:2404–12.
- Lorenz K, Bruhn G, Heumann C, Hoffmann T, Netuschil L. How to select study designs and parameters to investigate the effect of mouthrinses? Part II: comparisons between the parameters used. *J Physiol Pharmacol*. 2009a;60 Suppl 8:85–90.
- Lorenz K, Bruhn G, Netuschil L, Heumann C, Hoffmann T. How to select study designs and parameters to investigate the effect of mouthrinses? Part I: rationale and background. *J Physiol Pharmacol*. 2009b;60 Suppl 8:77–83.
- Lorenz K, Mayer D, Bruhn G, Noack B, Brex M, Heumann C, Toutenburg H, Netuschil L, Nagl M, Gottardi W, Hoffmann T. Effect of N-chlorotaurine mouth rinses on plaque regrowth and plaque vitality. *Clin Oral Investig*. 2009c;13:9–14.
- Mandel ID, Pihlstrom BL, Ramfjord S, et al. Periodontology and occlusion at Michigan. *J Dent Res*. 1997;76:1716–9.
- Mariath AA, Bressani AE, Haas AN, Araujo FB, Rösing CK. Professional flossing as a diagnostic method for gingivitis in the primary dentition. *Braz Oral Res*. 2008;22:316–21.
- Massler M. The P-M-A index for the assessment of gingivitis. *J Periodontol*. 1967;38:592–601.
- Matthijs S, Sabzevar MM, Adriaens PA. Intra-examiner reproducibility of 4 dental plaque indices. *J Clin Periodontol*. 2001;28:250–4.
- McBain AJ, Bartolo RG, Catrenich CE, Charbonneau D, Ledder RG, Gilbert P. Effects of a chlorhexidine gluconate-containing mouthwash on the vitality and antimicrobial susceptibility of in vitro oral bacterial ecosystems. *Appl Environ Microbiol*. 2003;69:4770–6.
- McCracken GI, Preshaw PM, Steen IN, Swan M, deJager M, Heasman PA. Measuring plaque in clinical trials: index or weight? *J Clin Periodontol*. 2006;33:172–6.
- Moustakis VS, Laine ML, Koumakis L, et al. Modeling genetic susceptibility: a case study in periodontitis. In: Combi C, Tucker A, editors. *Proceedings of IDAMAP-2007: intelligent data analysis in biomedicine and pharmacology*. Amsterdam: Artificial Intelligence in Medicine (AIME) Workshop; 2007. p. 59–64.
- Mühlemann HR. Psychological and chemical mediators of gingival health. *J Prev Dent*. 1977;4:6–17.
- Mühlemann HR, Son S. Gingival sulcus bleeding—a leading symptom in initial gingivitis. *Helv Odontol Acta*. 1971;15:107–13.
- Müller HP, Barrieshi-Nusair KM. Gingival bleeding on repeat probing after different time intervals in plaque-induced gingivitis. *Clin Oral Investig*. 2005;9:278–83.
- Müller HP, Heinecke A. The influence of gingival dimensions on bleeding upon probing in young adults with plaque-induced gingivitis. *Clin Oral Investig*. 2002;6:69–74.
- Nasi JH. Background to, and implementation of, the Periodontal Screening and Recording (PSR) procedure in the USA. *Int Dent J* 1994 44: 585–588.
- National Institute of Dental Research. U.S. Public Health Service, Oral Health of United States Adults; National Findings. Bethesda, MD: National Institute of Dental Research; 1987. NIH publication number 87-2868.
- Nesse W, Abbas F, van der Ploeg I, Spijkervet FKL, Dijkstra PU, Vissink A. Periodontal inflamed surface area: quantifying inflammatory burden. *J Clin Periodontol*. 2008;35:668–73.
- Nesse W, Linde A, Abbas F, Spijkervet FKL, Dijkstra PU, de Brabander EC, Gerstenbluth I, Vissink A. Dose-response relationship between periodontal inflamed surface area and HbA1c in type 2 diabetics. *J Clin Periodontol*. 2009;36:295–300.
- Netuschil L, Reich E, Brex M. Direct measurement of the bactericidal effect of chlorhexidine on human dental plaque. *J Clin Periodontol*. 1989;16:484–8.
- Netuschil L, Weiger R, Preisler R, Brex M. Plaque bacteria counts and vitality during chlorhexidine, meridol and listerine mouthrinses. *Eur J Oral Sci*. 1995;103:355–61.
- Netuschil L, von Ohle C, Brex M. Die vitalfluoreszenztechnik in der plaqueforschung. *Parodontologie*. 1996;7:293–305.
- Netuschil L, Reich E, Unteregger G, Sculean A, Brex M. A pilot study of confocal laser scanning microscopy for the assessment of undisturbed dental plaque vitality and topography. *Arch Oral Biol*. 1998;43:277–85.
- Newbrun E. Indices to measure gingival bleeding. *J Periodontol*. 1996;67:555–61.
- Newby EE, Bordas A, Kleber C, Milleman J, Milleman K, Keogh R, Murphy S, Butler A, Bosma ML. Quantification of gingival contour and volume from digital impressions as a novel method for assessing gingival health. *Int Dent J*. 2011;61 Suppl 3:4–12.
- Nowicki D, Vogel R, Melcer S, Deasy M. The gingival bleeding time index. *J Periodontol*. 1981;52:260–2.

- O'Leary T. The periodontal screening examination. *J Periodontol.* 1967;38(Suppl):617–24.
- O'Leary TJ, Drake RB, Naylor JE. The plaque control record. *J Periodontol.* 1972;43:38.
- Okada M, Kuwahara S, Kozai K, Kawamura M, Nagasaka N. The efficacy of an oral rating index for children for screening gingival health and oral hygiene status. *Pediatr Dent J.* 1999;9:91–7.
- Okada M, Kuwahara S, Kaihara Y, Ishidori H, Kawamura M, Miura K, Nagasaka N. Relationship between gingival health and dental caries in children aged 7–12 years. *J Oral Sci.* 2000;42:151–5.
- Okada M, Kawamura M, Kaihara Y, Matsuzaki Y, Kuwahara S, Ishidori H, Miura K. Influence of parents' oral health behaviour on oral health status of their school children: an exploratory study employing a causal modelling technique. *Int J Paediatr Dent.* 2002;12:101–8.
- Okada M, Kawamura M, Hayashi Y, Takase N, Kozai K. Simultaneous interrelationship between the oral health behavior and oral health status of mothers and their children. *J Oral Sci.* 2008;50:447–52.
- Persson RE, Kiyak AH, Wyatt CCI, MacEntee M, Persson GR. Smoking, a weak predictor of periodontitis in older adults. *J Clin Periodontol.* 2005;32:512–7.
- Petersen PE, Ogawa H. Strengthening the prevention of periodontal disease: the WHO approach. *J Periodontol.* 2005;76:2187–93.
- Pretty IA, Edgar WM, Higham SM. A study to assess the efficacy of a new detergent free, whitening dentifrice in vivo using QLF planimetric analysis. *Br Dent J.* 2004;197:561–6.
- Pretty IA, Edgar WM, Smith PW, Higham SM. Quantification of dental plaque in the research environment. *J Dent.* 2005;33:193–207.
- Quigley GA, Hein JW. Comparative cleansing efficiency of manual and power brushing. *J Am Dent Assoc.* 1962;65:26–9.
- Quirynen M, van Steenberghe D. Is early plaque growth rate constant with time? *J Clin Periodontol.* 1989;16:278–83.
- Quirynen M, Dekeyser C, van Steenberghe D. The influence of gingival inflammation, tooth type, and timing on the rate of plaque formation. *J Periodontol.* 1991;62:219–22.
- Ramfjord SP. Indices for prevalence and incidence of periodontal disease. *J Periodontol.* 1959;30:51–9.
- Ramfjord SP. The periodontal status of boys 11 to 17 years old in Bombay, India. *J Periodontol.* 1961;32:237–48.
- Renton-Harper P, Claydon N, Warren P, Newcombe RG, Addy M. Conversion of plaque-area measurements to plaque index scores. An assessment of variation and discriminatory power. *J Clin Periodontol.* 1999;26:429–33.
- Rosin M, Splieth C, Hesser M, et al. Quantification of gingival edema using a new 3-D laser scanning method. *J Clin Periodontol.* 2002;29:240–6.
- Rundegren J, Hvid EB, Johansson M, Åström M. Effect of 4 days of mouth rinsing with delmopinol or chlorhexidine on the vitality of plaque bacteria. *J Clin Periodontol.* 1992;19:322–5.
- Russell AL. A system of classification and scoring for prevalence surveys of periodontal disease. *J Dent Res.* 1956;35:350–9.
- Rustogi KN, Curtis JP, Volpe AR, Kemp JH, McCool JJ, Korn LR. Refinement of the modified navy plaque index to increase plaque scoring efficiency in gumline and interproximal tooth areas. *J Clin Dent.* 1992;3(Suppl C):C9–12.
- Sagel PA, Lapujade PG, Miller JM, Sunberg RJ. Objective quantification of plaque using digital image analysis. *Monogr Oral Sci.* 2000;17:130–43.
- Sandler HC, Stahl SS. Measurement of periodontal disease prevalence. *J Am Dent Assoc.* 1959;58:93–7.
- Saxer UP, Mühlemann HR. Motivation and education [Article in German]. *SSO Schweiz Monatsschr Zahnheilkd.* 1975;85:905–19.
- Saxton CA, van der Ouderaa FJ. The effect of a dentifrice containing zinc citrate and Triclosan on developing gingivitis. *J Periodontal Res.* 1989;24:75–80.
- Schick RA, Ash MM. Evaluation of the vertical method of toothbrushing. *J Periodontol.* 1961;32:346–53.
- Sculean A, Auschill TM, Donos N, Brex M, Arweiler NB. Effect of an enamel matrix protein derivative (EmdogainA) on ex vivo dental plaque vitality. *J Clin Periodontol.* 2001;28:1074–8.
- Seymour RA, Smith DG, Turnbull DN. The effects of phenytoin and sodium valproate on the periodontal health of adult epileptic patients. *J Clin Periodontol.* 1985;12(6):413–9.
- Shaloub A, Addy M. Evaluation of accuracy and variability of scoring-area-based plaque indices. A laboratory model. *J Clin Periodontol.* 2000;27:16–21.
- Shaw L, Murray JJ. A new index for measuring extrinsic stain in clinical trials. *Community Dent Oral Epidemiol.* 1977;5:116–20.
- Schour L, Massler M. Survey of gingival disease using the PMA index. *J Dent Res.* 1948;27:733.
- Silness J, Loe H. Periodontal disease in pregnancy. 1. Prevalence and severity. *Acta Odontol Scand.* 1963;21:533–51.
- Silness J, Loe H. Periodontal disease in pregnancy. 2. Correlation between oral hygiene and periodontal condition. *Acta Odontol Scand.* 1964;22:121–35.
- Smith RN, Brook AH, Elcock C. The quantification of dental plaque using an image analysis system: reliability and validation. *J Clin Periodontol.* 2001;28:1158–62.
- Smith RN, Rawlinson A, Lath D, Elcock C, Walsh TF, Brook AH. Quantification of dental plaque on lingual tooth surfaces using image analysis: reliability and validation. *J Clin Periodontol.* 2004;31:569–73.
- Smith RN, Rawlinson A, Lath DL, Brook AH. A digital SLR or intra-oral camera: preference for acquisition within an image analysis system for measurement of disclosed dental plaque area within clinical trials. *J Periodontal Res.* 2006;41:55–61.
- Smith RN, Lath DL, Rawlinson A, et al. Gingival inflammation assessment by image analysis: measurement and validation. *Int J Dent Hyg.* 2008;6:137–42.

- Söder P, Jin LJ, Söder B. Computerized planimetric method for clinical plaque measurement. *Scand J Dent Res*. 1993;101:21–5.
- Söder B, Jin LJ, Lundquist G, Söder PO. A longitudinal investigation of the individual consistency of plaque levels in adults. *Acta Odontol Scand*. 1995;53:72–4.
- Söderling E, Mäkinen KK, Chen CY, Pape Jr HR, Loesche W, Mäkinen PL. Effect of sorbitol, xylitol, and xylitol/sorbitol chewing gums on dental plaque. *Caries Res*. 1989;23:378–84.
- Staudt CB, Kinzel S, Hassfeld S, Stein W, Staehle HJ, Dörfer CE. Computer-based intraoral image analysis of the clinical plaque removing capacity of 3 manual toothbrushes. *J Clin Periodontol*. 2001;28:746–52.
- Suomi JD. Periodontal disease and oral hygiene in an institutionalized population: report of an epidemiological study. *J Periodontol*. 1969;40:5–10.
- Suomi JD, Barbano JP. Patterns of gingivitis. *J Periodontol*. 1968;39:71–4.
- Takahashi Y, Kamijyo H, Kawanishi S, Takaesu Y. Presence and absence of bleeding in association with calculus in segments given code 2 in the community periodontal index of treatment needs (CPITN). *Community Dent Oral Epidemiol*. 1988;16:109–11.
- Tezal M, Grossi SG, Ho AW, Genco RJ. Alcohol consumption and periodontal disease the third National Health and Nutrition Examination Survey. *J Clin Periodontol*. 2004;31:484–8.
- Tinanoff N, Hock J, Camosci D, Helldén L. Effect of stannous fluoride mouthrinse on dental plaque formation. *J Clin Periodontol*. 1980;7:232–41.
- Turesky S, Gilmore ND, Glickman I. Reduced plaque formation by the chloromethyl analogue of vitamin C. *J Periodontol*. 1970;41:41–3.
- van der Velden U. Probing force and the relationship of the probe tip to the periodontal tissues. *J Clin Periodontol*. 1979;6:106–14.
- Van der Velden U. Influence of probing force on the reproducibility of bleeding tendency measurements. *J Clin Periodontol*. 1980;7:421–7.
- Van der Weijden GA, Timmerman MF, Saxton CA, Russell JJ, Huntington E, Van der Velden U. Intra-/inter-examiner reproducibility study of gingival bleeding. *J Periodontol Res*. 1994a;29:236–41.
- Van der Weijden GA, Timmerman Nijboer A, Reijerse E, Van der Velden U. Comparison of different approaches to assess bleeding on probing as indicators of gingivitis. *J Clin Periodontol*. 1994b;21:589–94.
- Vettore MV, Lamarca Gde A, Leão AT, Sheiham A, Leal Mdo C. Partial recording protocols for periodontal disease assessment in epidemiological surveys. *Cad Saude Publica*. 2007;23:33–42.
- Volpe AR, Manhold JH. A method of evaluating the effectiveness of potential calculus inhibiting agents. *N Y State Dent J*. 1962;28:289–90.
- von Ohler C, Weiger R, Decker E, Schlagenhauf U, Brex M. The efficacy of a single pocket irrigation on subgingival microbial vitality. *Clin Oral Investig*. 1998;2:84–90.
- Weiger R, Netuschil L, Brex M. Relationship between bacterial counts, microbial vitality and the accumulation of supragingival dental plaque in humans. *J Periodontol Res*. 1992;27:575–80.
- Weiger R, Netuschil L, Von Ohle C, Brex M. Microbial vitality of supragingival dental plaque during initial stages of experimental gingivitis in humans. *J Periodontol Res*. 1995;30:204–9.
- Weiger R, Netuschil L, Wester-Ebbinghaus T, Brex M. An approach to differentiate between antibacterial and antiadhesive effects of mouthrinses in vivo. *Arch Oral Biol*. 1998;43:559–65.
- World Health Organization. Oral health surveys: basic methods. Geneva: World Health Organization; 1997.
- Yeganeh S, Lynch E, Jovanovski V, et al. Quantification of root surface plaque using a new 3-D laser scanning method. *J Clin Periodontol*. 1999;26:692–7.
- Zaura-Arite E, van Marle J, ten Cate JM. Confocal microscopy study of undisturbed and chlorhexidine-treated dental biofilm. *J Dent Res*. 2001;80:1436–40.

19.1 Introduction

Epidemiology can be defined as ‘the study of the distribution and determinants of disease in human populations’. Distribution refers to the frequency and the severity of the disease within a population or its subgroups. Frequency is measured in terms of prevalence and incidence and severity or extent by indexes or scales that quantify the levels of the disease. The determinants of disease refer to factors that increase the risk of disease onset or that identify population subgroups most likely to have the disease under consideration (Locker et al. 1998).

In periodontal epidemiology, a useful distinction can be made between descriptive and analytical studies. Descriptive studies aim to describe periodontal disease in a population—who has it, to what extent do they have it, and how much they have—while analytical studies seek to explain the occurrence of disease by examining its association with putative risk factors (Thomson and Williams 2002).

19.2 Epidemiological Study Designs

The epidemiology of a disease or condition includes the prevalence of the disease (how much is out there), the incidence of the disease (how much new disease can we expect in a given time period), and the characteristics of the people at risk for the disease. This information is generated by three types of observational population-based

studies: cohort, case–control, and cross-sectional studies (Beck and Löe 1993).

19.2.1 Cohort Studies

Cohort studies are also called follow-up or incidence studies because they are prospective in design. Usually, a disease-free population (preferably representative) for which the explanatory variable(s) of interest are known is followed over time to see who develops periodontal disease. Because this type of observational study most closely resembles an experimental design (and it is known that the explanatory variable(s) existed in time before studying the periodontal disease), cohort studies are often used to make causal inferences and to identify risk factors (Beck and Löe 1993).

There are important conceptual ties between randomized and non-randomized cohorts and between non-randomized cohorts and case–control studies. If the assignment mechanism for exposure is not random, the reference cohort can be thought of as a ‘confounded’ substitute for the counterfactual experience. Cohort studies are inherently ‘prospective’, in that exposure levels of individuals in the cohort are measured at a time prior to any disease outcomes of interest. Individuals are followed over time, and person–time is accrued until development (incidence) of the outcome, death, or loss to follow-up. The accumulation of person–time allows direct calculation of the incidence rate, the most preferred

measure of disease frequency for studying causal effects. The source population enumerated in a cohort study is most often formed on the basis of some common characteristic or experience. For general cohorts, individuals are separated according to their level of the exposure variable at baseline and often at subsequent intervals over a specified follow-up period as well (time-varying exposure) (Heaton and Dietrich 2012).

The usual measures of association include the risk ratio, relative risk, and attributable risk. This design may be useful for following the natural course or progression of periodontal disease in a group of subjects when little or no control and monitoring are imposed (Kingman and Albandar 2002).

19.2.2 Case-Control Studies

Of all analytic designs, case-control studies are probably the most misunderstood. Foremost among such misconceptions is the view that case-control studies are inferior to cohort studies in terms of producing valid effect estimates. This and other common misconceptions are probably the result of what Fienstein terms the ‘**trohoc fallacy**’—the fallacious thinking that case-control studies are simply cohort studies performed backwards and that they are therefore inherently ‘retrospective’ (Heaton and Dietrich 2012).

Case-control studies are retrospective in design and provide a method for investigating the factors that may prevent or cause the disease. The method involves the comparison of patients (cases) with a group of control. The comparison is aimed at discovering factors that may differ in the two groups and that explain the occurrence of the disease in the cases. Case-control studies are often used when the disease occurs infrequently and when one wants to generate ideas to test in cross-sectional and cohort studies (Beck and Löe 1993).

Assuming that the underlying cohort is fixed and cases represent incident disease, if controls are selected from among all subjects at baseline or the ‘at risk’ population, they will represent the distribution of exposure among the total population at risk at study baseline (denominator of the risk),

and the observed odds ratio will therefore approximate the risk ratio with no need for a rare disease assumption. This particular approach is sometimes referred to as a **case-cohort design** (Heaton and Dietrich 2012).

The usual measure of association is the odds ratio. A case-control study can provide important information related to associations between exposure levels for rare diseases but are not capable of providing estimates of disease occurrence. A major concern in a case-control study is the selection of appropriate controls: the healthy controls, case-control criteria, and the exposure frequencies among cases and controls (Kingman and Albandar 2002).

Although case-control studies and cohort studies are similar from a validity standpoint, they are often in direct opposition from a utility standpoint. Case-control studies may be the only useful alternative to a cohort study when diseases are rare, latency periods are long, and/or a study base has not previously been enumerated. Similarly, cohort studies are the only practical approach when the exposure of interest is sufficiently rare. By extension, cohort studies are particularly useful for studying multiple outcomes of a single exposure, and case-control studies are particularly useful for studying multiple exposures of a single outcome (Heaton and Dietrich 2012).

19.2.3 Cross-Sectional Studies

Cross-sectional studies usually study representative populations at one point in time to determine the prevalence of the disease at a certain moment and the population characteristics that are associated with having the disease. Prevalent data from cross-sectional studies serve a number of purposes: to monitor the amount of disease existing in a population, to delineate the characteristics of people who have the disease, to generate hypotheses regarding the aetiology of the disease, to plan staffing needs for preventive and treatment services, and to plan teaching programmes. A series of properly conducted prevalence studies over time can be used to monitor the trends of the disease, the effectiveness of preventive and treatment

programmes, and under certain conditions, to estimate the incidence rates of the disease (Beck and Löe 1993). Cross-sectional studies, however, can never distinguish between incidence and natural history of disease, regardless of variations on the design. Further, cross-sectional studies cannot provide effect estimates that approximate the risk or rate ratio, despite misconceptions perpetuated in the literature (Heaton and Dietrich 2012).

They are particularly useful for estimating prevalence of specific health and disease characteristics for large population groups or specific subgroups. An example might involve a risk assessment of factors, such as behavioural traits, genetic and environmental, and other confounding factors thought to be associated with disease, health, or quality of life. Large health surveys can provide important information for less common phenotypes and genotypes or for specific subgroups that would normally not be adequately represented in smaller studies (Kingman and Albandar 2002).

19.3 Issues of Study Design

19.3.1 Methods for Sampling Subjects

There are several different sampling techniques available.

19.3.1.1 Simple Random Sampling

In this case, each individual is chosen entirely by chance, and each member of the population has an equal chance, or probability, of being selected. One way of obtaining a random sample is to give each individual in a population a number and then use a table of random numbers to decide which individuals to include (Barratt 2009).

19.3.1.2 Systematic Random Sampling

Individuals are selected at regular intervals from a list of the whole population. The intervals are chosen to ensure an adequate sample size. For example, every 10th member of the population is included. This is often convenient and easy to use, although it may also lead to bias for reasons outlined below (Barratt 2009).

Hugoson et al. (2008) investigated the prevalence and distribution of gingivitis and periodontitis in a Swedish population over the 30 years 1973–2003 using a systematic random sampling. In 1973, a random sample of individuals aged 3, 5, 10, 15, 20, 30, 40, 50, 60, and 70 years from four parishes were examined clinically and radiographically. In addition to the clinical examinations, subjects were asked about dental care habits and knowledge of oral health. The sample was chosen based on date of birth (between March and May), and all subjects in each age group were listed in chronological order so that each age group had 140–170 individuals. The first individual on each age group list was invited to the examination. In the event of non-attendance, the next individual on the list was invited until 100 participants in each age group had been chosen. In 1983, 1993, and 2003, new random samples of subjects were selected, consisted of 130 individuals in each of the age groups used in 1973 plus a group aged 80 years. Of the selected cases, 677, 655, and 589 individuals 20–80 years old showed up at the three different examinations respectively.

19.3.1.3 Stratified Random Sampling

In this method, the population is first divided into subgroups (or strata) who all share a similar characteristic. It is used when we might reasonably expect the measurement of interest to vary between the different subgroups. Gender or smoking habits would be examples of strata. The study sample is then obtained by taking samples from each stratum. In a stratified sample, the probability of an individual being included varies according to known characteristics, such as gender, and the aim is to ensure that all subgroups of the population that might be of relevance to the study are adequately represented. The fact that the sample was stratified should be taken into account at the analysis stage (Barratt 2009).

For example, Baharin et al. (2006) tested the hypothesis that smokers have more disease in the upper anterior region using a retrospective stratified random sample of 49 non-smokers, and 39 heavy smokers (≥ 20 cigarettes/day) was obtained from a total of 3,678 referred patients with adult periodontitis. Inclusion criteria were

(1) Patients diagnosed with moderate-to-severe periodontitis, (2) aged between 46 and 60 years, (3) patient with no significant medical history, (4) smoking patients reported smoking ≥ 20 cigarettes/day, and (5) Non-smoking patients reported never smoking. Exclusion criteria were (1) patients with abnormal tooth and root form; (2) patients with a significant medical history, such as diabetes or drugs, affecting inflammation; and (3) past smokers. A random sample, stratified for age, was selected from the 650 records that satisfied these criteria. The selection was achieved by dividing the records into 3-year age bands and using a random number generation program (Stata Co., College Station, TX, USA). One hundred and twenty records were generated, and 88 records (smoker=39, non-smoker=49) were obtained with dental panoramic tomographs (DPTs) that allowed assessment of bone levels in all regions of the mouth. The mean age of the subjects was similar between smokers (52.2 years, standard deviation (SD) ± 4.4) and non-smokers (52.9 ± 4.2 years) (Baharin et al. 2006).

19.3.1.4 Cluster Sampling

In a clustered sample, subgroups of the population are used as the sampling unit, rather than individuals. The population is divided into subgroups, known as clusters, and a selection of these are randomly selected to be included in the study. All members of the cluster are then included in the study. Clustering should be taken into account in the analysis (Barratt 2009).

19.3.1.5 Convenience Sampling

Convenience sampling is perhaps the easiest method of sampling because participants are selected in the most convenient way and are often allowed to choose or volunteer to take part. Good results can be obtained, but the data set may be seriously biased because those who volunteer to take part may be different from those who choose not to (Barratt 2009).

19.3.1.6 Snowball Sampling

This method is commonly used in social sciences when investigating hard to reach groups. Existing subjects are asked to nominate further subjects

known to them, so the sample increases in size like a rolling snowball (Barratt 2009).

19.3.1.7 Multistage Sampling

These sampling designs involve sampling in stages. The sampling units at the 1st stage are termed primary sampling units (PSUs). Multistage sampling designs often involve many stages, with different basic sampling strategies employed at each stage (Kingman and Albandar 2002).

Susin et al. (2004) estimated the percentages of cases with severe periodontal attachment loss attributable to cigarette smoking in a representative adult urban population in southern Brazil. A representative sample comprising 853 dentate individuals (age 30–103 years) was selected by a multistage, probability sampling method from a larger sample representative of subjects aged 14 years and older among the population of Porto Alegre. Using area maps, the Porto Alegre metropolitan area was divided into 90 geographic areas 10 km² each. Using the 1991 census data (Brazilian Institute of Geography and Statistics IBGE 1991) and other relevant municipal information (Rio Grande do Sul state government agency for metropolitan affairs METROPLAN 1997), these geographic areas were stratified into 13 (14.4%) high-income and 77 (85.6%) low-income status areas. Low-income geographic areas were defined as areas in which more than 40% of the head of the households had a monthly income of 42 standard Brazilian salaries (about US\$ 180), and high-income areas were those with a higher level of income. Within each of these two income strata, primary sampling units (PSUs) were selected randomly with a probability proportional to size and using a sampling frame of these PSUs. A total of 11 geographic areas were selected and included two (18.2%) areas with high and nine (81.8%) areas with low-income status (Susin et al. 2004).

The second stage consisted of selecting area sectors within each geographic area. The area sectors have been defined by IBGE as map areas comprising approximately 300 households each. The sectors were selected randomly within each geographic area, and the number of sectors

selected was proportional to the number of sectors in each area. Thirty (3.5%) sectors were selected, out of a total of 846 eligible sectors. Approvals for conducting the study were sought separately in each sector from key community, religious, and/or administrative leaders. Permission and/or support were granted to access 29 of these sectors, whereas permission to access one sector was denied (Susin et al. 2004).

The third stage included selecting households within each of the 29 sectors. It was estimated that approximately 25 households were needed per sector to provide a sufficient number of subjects in the sample. In each sector, a starting point for the selection of households was established on area maps and was provided independently by the IBGE. Households were sampled consecutively beginning with the next block after the starting point and until the preset number of households was reached (Susin et al. 2004).

Consenting household members who were 14 years of age or older were examined, and subjects 30 years or older were included in this study. Exclusion criteria were presence of diseases/conditions that may pose health risks to the participant or examiner or that may interfere with the clinical examination. Hence, subjects were excluded if they were diagnosed with psychiatric problems or intoxicated with alcohol or drugs. Individuals requiring a prophylactic regimen of antibiotics were provided with the appropriate medicine before the clinical examination (Susin et al. 2004).

19.3.1.8 Bias in Sampling

There are five important potential sources of bias that should be considered when selecting a sample, by whatever method (Barratt 2009):

- (a) Any changes from the pre-agreed sampling rules can introduce bias.
- (b) Bias is introduced if people in hard to reach groups are omitted.
- (c) Replacing selected individuals with others, for example, if they are difficult to contact, also introduces bias.
- (d) It is important to try and maximize the response rate to a survey; low response rates can introduce bias.

- (e) If an out of date list is used as the sample frame, it may also introduce bias, if it excludes people who have recently moved to an area, for example.

19.3.1.9 Sampling Weights

Weighted estimates of population parameters can be derived by several software packages. However, a major problem one must confront when reporting survey data is the derivation of the appropriate variances and standard errors of the weighted disease estimates. For a complex multistage probability sample, these can be tedious and complicated. The statistical methodology and available software needed for performing these analyses from health surveys has expanded greatly in the past few decades. Major developments would include refined methods for use in multistage sampling, efficient techniques for oversampling specific subpopulations, and methods for estimating variances in complex survey designs. The survey design and analysis package (SUDAAN) is a multifaceted software package that handles correlated data structure from complex survey designs (Kingman and Albandar 2002).

19.3.2 Methods for Sampling Sites Within Subjects (Partial Recording Protocols)

Most survey methods use a full-mouth assessment of periodontal diseases. This involves the examination of 6 sites on all teeth, involving up to 168 sites per mouth (excluding 3rd molars). This results in 168 measurements for each periodontal pocket depth and clinical attachment level parameter, involving at least 336 pieces of clinical data for each subject. Because of the time, logistic, and cost constraints for a full-mouth assessment in epidemiological surveys involving large population samples, the full-mouth clinical assessment of periodontal diseases is impractical. Therefore, **partial recording protocols** have been developed and used to characterize the periodontal status of subjects and populations (Kingman and Albandar 2002).

Symmetry or the intrinsic property of an object to remain invariant under certain classes of transformations (such as reflection, rotation, inversion, or more abstract mathematical operations) is a common feature of biological systems and can be observed in multiple aspects of the human body. An object is amphichiral (also called re-flexible) if it is superposable with its image in a plane mirror (Mombelli and Meier 2001). Human periodontal disease has been described as a symmetrical disorder (Mombelli and Meier 2001). Analysis of the incidence of active disease and bacterial colonization at symmetrical sites in humans has concluded that the development of disease in the light of existing contralateral lesions is not solely dependent on the presence of local factors (Mombelli and Meier 2001). In an experimental model of periodontitis in the rat, potential symmetry of disease was also described in experimental animals. An alternative explanation for these results is that contralateral inflammation with resulting bone loss is stimulated through a neurogenic mechanism (Abd El-Aleem et al. 2004; Dumitrescu et al. 2004).

Irrespective of the methods selected for assessing periodontal disease prevalence—clinical or radiographic—the majority of epidemiological surveys have employed partial recordings. The rationale for the use of partial examinations has been the assumption that (1) the time required for the performance of a partial survey, and consequently its cost, is significantly decreased and (2) provided that the examined segments adequately reflect the periodontal condition of the entire dentition, the amount of information lost is kept to a minimum (Papapanou et al. 1993).

Different models of partial-mouth protocols such as reduction in quadrants, teeth, sites, or combinations of the above have been suggested:

19.3.2.1 Ramfjord Teeth

One partial recording system was proposed by Ramfjord (1959) and was based on the examination of **6 index teeth (16, 22, 24, 36, 42, and 44)** (Papapanou et al. 1993). It has been showed that they produce systematic underestimation of disease prevalence, especially for the more severe

forms of the disease, but are adequate for gingivitis status (Fleiss et al. 1987).

19.3.2.2 CPITN Teeth

Examination of a subset of **10 index teeth (17, 16, 11, 26, 27, 37, 36, 31, 46, 47)** has been recommended by the CPITN use in epidemiological studies of adults (Ainamo et al. 1982; Papapanou et al. 1993). When the results of a full-mouth examination were compared with the results of examining only the CPITN selection of 10 index teeth 17/16, 11, 26/27, 47/46, 31, and 36/37 for estimates of prevalence and severity of the conditions assessed with the CPITN, it has been showed that the prevalence of pockets, whether moderate or deep, was underestimated in virtually all age groups (Baelum et al. 1993).

19.3.2.3 Random Half-Mouth

Other commonly employed partial recording methods are based on assessments derived from **one upper and one lower quadrant** (Carlos et al. 1986). These should be randomly selected and may belong to the same half or to diagonally opposite halves of the dentition (Papapanou et al. 1993). The use of half-mouth examinations is based on the assumption that oral conditions are distributed bilaterally but vary among the teeth on each side in a reasonably symmetrical pattern (Hunt 1987).

19.3.2.4 Fixed Sites Full-Mouth

These partial recording protocols are characterized by evaluating a fixed set of sites on all 28 teeth in the dentition (excluding 3rd molars). For example, the full-mouth versions of the National Institute of Dental and Craniofacial Research (NIDCR) partial recording protocols consist of all MB–B (mesio-buccal and mid-buccal) and all MB–B–DB sites. Any set of fixed sites per tooth for all teeth would comprise a fixed-site full-mouth protocol (Kingman and Albandar 2002).

Kingman et al. (1988) used three large data sets containing full-mouth examinations (mesio-buccal, mid-buccal, disto-buccal, and mid-lingual (MBDL) sites per tooth) for either attachment loss or probing pocket depths to investigate the magnitudes of systemic error that occur by

employing four specific partial-mouth scores (M, MB, MBD, and MBDL sites per tooth), which are based on the random half-mouth technique. For prevalence of disease, the sensitivity of a partial score was a function of the disease level in the population. All four partial scores were sensitive enough to adequately portray true prevalence using the 2-mm demarcation value, the MBD and MBDL scores might still be considered adequate for the 4-mm value, but none of these partial scores were adequate for the 7-mm value. For disease severity, the MBDL score produced unbiased estimates, the others were biased. The relative biases for the MB and MBD scores were all under 10% (in absolute value), but the M score produced severe relative biases, 24% for probing pocket depths and -12% for attachment loss (Kingman et al. 1988).

19.3.2.5 Random Subsets of Teeth (or Sites)

This method consists of the selection of a random subset of teeth (or a random or fixed subset of sites per tooth). For example, one could randomly select 28 sites per mouth for the subjects participating in the survey (Kingman and Albandar 2002).

One method that subsamples' both teeth and sites system was developed by the National Institute of Dental Research's (NIDCR) to conduct national surveys in situations where time and financial resources were limited. The NIDCR method employed a strategy of randomly selecting an upper and a contralateral lower quadrant in which to conduct the periodontal measures. In addition, they examined only the mesio-buccal and buccal sites on all teeth in those quadrants. As Beck et al. (2006) summarized, this system has been used to conduct several large studies, such as a study of employed and retired adults (Miller et al. 1987) and the National Health and Nutrition Examination Survey (NHANES III) (Albandar et al. 1999). In addition, the method has been used by others conducting smaller studies (Beck et al. 1995; Slade et al. 1993).

Recently, Timmerman et al. (2002) reported the inaccuracy for the 4-site data (deepest pocket per quadrant that showed the most attachment

loss and bleeding on probing) when used as a descriptive for changes in the periodontal condition on a full-mouth level over a 7-year period.

19.3.2.6 Comparison of Disease Estimates by Types of Partial Recording Protocol

Different models of partial-mouth protocols such as reduction in quadrants, teeth, sites, or combinations of the above have been compared with full-mouth examinations (Hunt and Fann 1991; Hunt 1987; Agerholm and Ashley 1996; Ainamo and Ainamo 1985; Beck and Löe 1993; Downer 1972; Dowsett et al. 2002; Eaton et al. 2001; Fleiss et al. 1987; Kingman and Albandar 1997; Kingman and Albandar 2002; Kingman et al. 1988, 2008; Mills et al. 1975; Mumghamba et al. 2004; Owens et al. 2003; Papapanou et al. 1993; Silness and Røynstrand 1988; Thomson and Williams 2002; Mumghamba et al. 2004) (Table 19.1).

As Albandar (2011) summarized, it is well documented that partial-mouth protocols, which are often used in surveys, do underestimate in varying degrees the prevalence of periodontal diseases (Kingman and Albandar 2002; Susin et al. 2005; Kingman et al. 2008; Eaton et al. 2001; Albandar 2011). Furthermore, different partial-mouth protocols (using different combinations of sites) show different degrees of underestimation of these diseases. Currently, there is no consensus on which 'representative' sites to use for a partial-mouth examination (Albandar 2011). It has been suggested that the partial recording protocol should include the same proportion (two-thirds) of proximal sites as the full-mouth protocol so that one can minimize the possible bias in estimating average levels of periodontal attachment loss (AL) or pocketing (Kingman et al. 2008). However, when estimating the severity of periodontitis, the choice of specific sites is probably more important than the number of sites. Notably, other factors may influence the amount of underestimation when using partial-mouth protocols, such as the prevalence and severity of periodontal disease and the extent of missing teeth in the population under study (Albandar 2011). Hence, the bias caused by a given system may increase with the increase

Table 19.1 Summary of recent selected studies evaluating the partial recording protocols for periodontal disease assessment in epidemiological surveys

Author, year	Study population	Probe used	Partial-mouth protocols	Main findings
Eaton et al. (2001)	100 subjects aged 16–20 years (mean 17 years)	Williams probe	Ramfjord index teeth (3, 9, 12, 19, 25, 28) and the Periodontal Index for Treatment (PIT) teeth (3, 8, 14, 19, 24, 30)	The prevalence of lifetime cumulative attachment loss (LCAL) ≥ 1 mm was similar (approaching 100%) for the full-mouth and both partial-mouth recordings. However, as LCAL increased from a minimum of 1–3 mm, partial-mouth recording resulted in an underestimation of the prevalence of disease. LCAL ≥ 2 mm was underestimated by up to 22% and LCAL ≥ 3 mm by up to 36%. The extent of LCAL was less affected by partial-mouth recording, in that the percentage of sites with no sign of early attachment loss was underestimated by up to 11%. However, the percentage of sites with LCAL ≥ 1 mm and 2 mm were overestimated by 11% and 7%, respectively
Kingman and Albandar (2002)	The 266 study subjects were selected from the 13–17-year-olds originally examined during the 1986–1987 NIDR Children’s survey	NIDCR periodontal probe	(1) Ramfjord teeth, (2) CPITN teeth, (3) MB–B (mesio-buccal and mid-buccal sites per tooth) and MB–B–DB (mesio-buccal, mid-buccal, and disto-buccal) random half-mouth protocols, and the (4) MB–B and MB–B–DB full-mouth protocols. The MB–B–DB random half-mouth and full-mouth protocols are included so that the proximal-smooth site types are proportionally represented relative to the full-mouth 6-site per tooth score	Thus, the MB–B random half-mouth protocol correctly identifies 60% of the subjects with attachment loss ≥ 3 mm, whereas the MB–B–DB random half-mouth protocol correctly identifies 73%. In contrast, the MB–B full-mouth protocol correctly identifies 74%. Consequently, one does just as well estimating disease (prevalence of attachment loss ≥ 3 mm) using the MB–B–DB random half-mouth protocol as one does with the MB–B–full-mouth protocol
Dowsett et al. (2002)	292 subjects aged 18–78 years old	Not given	Half-mouth (random diagonal quadrants) and Ramfjord teeth assessment	For mean PI, GI, PD, and CAL, both half-mouth and Ramfjord teeth assessment provided an acceptable alternative to whole-mouth assessment (ICCs >0.92). For percentage of sites above a specified threshold, ICCs were generally greater than 0.90 in all age cohorts for half-mouth assessment but consistently lower for Ramfjord teeth assessment. Ramfjord teeth assessment also considerably underestimated disease prevalence compared with half-mouth assessment

Timmerman et al. (2002)	158 young subjects 15–25 years of age at baseline	4-site data (deepest approximal pocket in each quadrant that showed the most attachment loss and bleeding on probing)	For disease progression between baseline and 7-year follow-up, significant correlation coefficients were observed between the 4-site and full-mouth mean changes (pocket depth (PD): 0.80; attachment loss (AL): 0.70; plaque index (PI): 0.77). Regression coefficients were 0.51 for PD, 0.35 for AL, and 0.55 for PI. The precision of the estimate for the full-mouth mean, as predicted by the 4-site mean, is determined by the residual standard deviation. This was for PD 0.31 mm, for AL 0.31 mm, and for PI 0.29. Compared to the between-patient standard deviation of the full-mouth means, the residual standard deviations were high and illustrate the inaccuracy for the 4-site data when used as a descriptive for changes in the periodontal condition on a full-mouth level
Thomson and Williams (2002)	200 participants 25- and 26-year-olds	Periodontal examination at three sites per tooth in all four quadrants for gingival recession, probing depth, and CAL. The half-mouth analyses took three forms: (1) estimates from each of the left and right sides were obtained and compared, (2) estimates were obtained separately and compared for quadrants 1 and 3 (upper right and lower left) and quadrants 2 and 4 (upper left and lower right), and (3) estimates were obtained from a diagonal half-mouth count, whereby quadrants 1 and 3 were analysed for study participants whose identification number was odd, and quadrants 2 and 4 were analysed for the remainder	The difference in prevalence estimates obtained from the different methods was considerably greater for GR than for PD and CAL. The unadjusted odds ratio (OR) for the prevalence of 1 or more teeth with > or = mm of CAL among smokers was 2.3 (95% confidence interval [CI] 1.0, 5.3) using the full data set and 2.4 (95% CI 0.9, 6.1) using the diagonal half-mouth approach

(continued)

Table 19.1 (continued)

Author, year	Study population	Probe used	Partial-mouth protocols	Main findings
Owens et al. (2003)	108 adult subjects (≥18 years old)		• Six sites per tooth on all teeth examined in random diagonal quadrants (RD6) • Six sites per tooth on all teeth examined in upper and lower right quadrants (UR/LR) • Six sites per tooth on all teeth examined in upper and lower left quadrants (UL/LL) • Six sites per tooth on the Ramfjord teeth (Ramfjord) • Two sites per tooth on all teeth examined (WM2) • Two sites (MB and BB) per tooth on all teeth examined in random diagonal quadrants (RD2)	For mean PD, CAL, and REC, ICCs for all partial-mouth assessments were generally greater than 0.85, apart from those for mean PD where assessment was limited to two sites per tooth. In the case of disease extent and severity, half-mouth assessment of six sites per tooth provided the closest estimate of disease level to that reported for whole-mouth assessment. For half-mouth assessment, UR/LR assessment tended to result in an overestimation of disease levels, while random diagonal (RD6) and UL/LL underestimated disease extent and severity. For assessment of six sites per tooth in one upper and one lower quadrant, ICCs were consistently >0.80. Assessment of two sites per tooth or only Ramfjord teeth generally underestimated disease extent and severity and prevalence, compared to half-mouth assessment
Mumghamba et al. (2004)	192 patients aged from 15 to 77 years (36 years old on average)	William's probe	Ramfjord teeth (teeth numbers 16, 21, 24, 36, 41, 44). If a Ramfjord tooth was missing, a substitute tooth was selected as suggested by Fleiss et al. (1987) (teeth numbers 17, 11, 25, 37, 31, 45)	The correlation between the mean pocket depth calculated from the full-mouth and Ramfjord teeth was 0.96. The <i>b</i> coefficient for the mean pocket depth measured by Ramfjord teeth to predict the full-mouth mean was 0.94 and was not affected by adjustment for age, missing teeth, or sex

Susin et al. (2005)	1,460 dentate persons 14–103 years old	colour-coded probe graded at 1, 2, 3, 5, 7, 8, 9, and 10 mm	Seven PRP were assessed in this study: • Mesio-buccal and mid-buccal measurements on all teeth (MB–B, full-mouth) • Mesio-buccal, mid-buccal, and disto-buccal measurements on all teeth (MB–B–DB, full-mouth) • Mesio-buccal, mid-buccal, and disto-lingual measurements on all teeth (MB–B–DL, full-mouth) • Mesio-buccal and mid-buccal measurements on all teeth in one maxillary and one mandibular randomly selected quadrant (MB–B, half-mouth) • Mesio-buccal, mid-buccal, and disto-buccal measurements on all teeth in one maxillary and one mandibular randomly selected quadrant (MB–B–DB, half-mouth) • Mesio-buccal, mid-buccal, and disto-lingual measurements on all teeth in one maxillary and one mandibular randomly selected quadrant (MB–B–DL, half-mouth) • Mesio-buccal, mid-buccal, disto-buccal, disto-lingual, mid-lingual, and mesio-lingual sites on all teeth in one maxillary and one mandibular randomly selected quadrant (six sites, half-mouth)	For protocols that used only two or three sites, the system that used MB, B, and DL yielded the largest prevalence rates for both the half-mouth and full-mouth design. In most PRP, the bias in the prevalence estimates was highest for moderate (4–7 mm) attachment loss. A PRP based on the MB, B, and DL full-mouth performed the best among the seven PRP evaluated uniformly across the severity of attachment loss used to define disease. Generally, better sensitivities were seen in systems that used full-mouth than half-mouth design and in three site systems that used DL rather than DB sites. The three-site PRP incorporating the DL site produced less bias than the three-site PRP including the disto-buccal (DB) site. There was a 3–12% gain in sensitivity for 2- to 5-mm attachment loss thresholds for the three-site half-mouth PRP compared with the two-site MB and B half-mouth PRP
Beck et al. (2006)	15,792 community-dwelling residents aged 45–64 years		Three fixed-site selection methods (FSSMs): • Ramfjord method • Mesio-buccal and buccal sites for all teeth in one upper and the contralateral lower quadrant • Mesio-buccal, buccal, and disto-buccal sites for all teeth in one upper and the contralateral lower quadrant • Seven random-site selection methods (RSSMs) created by sampling the following number of sites: 84, 42, 36, 28, 20, 15, 10, and 6	Estimates of means, standard deviations (SD), relative bias, and relative error for RSSMs were almost identical to the full-mouth examination, but SDs increase slightly when fewer than 28 sites were sampled and relative bias and error increase for methods sampling fewer than 20 sites. The FSSMs had a very low relative error but much higher relative bias indicating underestimation. The FSSM with the smallest bias and error was the Ramfjord method, but the Random 36 method had less bias and less relative error. Both FSSMs and RSSMs underestimated prevalence, especially prevalence of less frequently occurring conditions, but most RSSMs were less likely to underestimate prevalence than the FSSMs

(continued)

Table 19.1 (continued)

Author, year	Study population	Probe used	Partial-mouth protocols	Main findings
Vettore et al. (2007)	156 subjects (over 30 years of age and have at least 15 teeth)	North Carolina periodontal probe	- Model I: all sites per tooth in the random half-mouth protocol randomly selecting one maxillary and mandibular quadrant - Model II: buccal sites in a full-mouth protocol - Model III: buccal sites in the random half-mouth protocol randomly selecting one maxillary and mandibular quadrant - Model IV: all sites per tooth using Community Periodontal Index teeth	In comparison with full-mouth examination, Model I did not show significant differences for periodontal pocket depth and clinical attachment level parameters. Compared with full-mouth examination, Model II and Model III underestimated the mean periodontal pocket depth by 1.54% and 1.15%, respectively. Model IV overestimated the mean periodontal pocket depth and clinical attachment loss by 12.7% and 14.83%. It was concluded that Model I appears to be adequate to substitute for the full-mouth examination to assess the prevalence and severity of chronic periodontal disease in adults
Kingman et al. (2008)	1,437 dentate Brazilian subjects 14–103 years old		Twenty different PRPs: - Twelve 1-site PRPs: six random half-mouth PRPs and six full-mouth PRPs. The single sites were MB, B, DB, ML, L, and DL - Two 2-site PRPs: one random half-mouth PRP (as in NHANES III) and one full-mouth PRP. The two sites were MB–B - Four three-site PRPs: two random half-mouth PRPs and two full-mouth PRPs. The two combinations of three sites were MB–B–DB (as in NHANES IV) and MB–B–DL - One random half-mouth six-site PRP. The six sites were MB–D–DB–ML–L–DL - One six-site PRP using the Ramfjord teeth. The six sites were MB–D–DB–ML–L–DL	Slightly higher levels of disease were evidenced on lingual than on buccal sites. Seven multi-site PRPs and the Ramfjord PRP produced small biases in mean probing pocket depth (MPPD) (–0.17 to 0.04 mm) and mean clinical attachment loss (MCAL) with relative biases under 8% and 4% in absolute value for MPPD and MCAL, respectively. Biases for full- and random half-mouth-based PRPs were similar. The three-site random half-mouth MB–B–DL and the Ramfjord PRPs produced the smallest biases, with relative biases <3% in absolute value for MPPD and MCAL.
Peres et al. (2011)	339 adolescents (1,993 Pelotas birth cohort at 12-year-old); 720 adults (1,982 Pelotas birth cohort study at 24 years old)	Not given	Four partial protocols were used: Ramfjord teeth (RT teeth numbers 3, 8, 12, 19, 24, and 28), Community Periodontal Index (CPI uses teeth numbers 2, 3, 8, 14, 15, 18, 19, 24, 30), and two random diagonal quadrants (1 and 3, and 2 and 4)	Two diagonal quadrants showed better accuracy, RT had the worst, while CPI presented an intermediate pattern when compared with full-mouth examination. For bleeding assessment in adolescence, RT and CPI underestimated by 18.4% and 16.2%, respectively, the true outcome prevalence, while among young adults, all partial protocols underestimated the prevalence. All partial protocols presented similar magnitude of association measures for all investigated periodontal potential risk factors

ICC Intra-class correlation coefficients

in the severity of AL within the population (Kingman and Albandar 2002; Albandar 2011).

In a recent study, Kingman et al. (2008) assessed bias magnitudes of periodontal disease severity estimates for specific partial recording protocols (PRPs) in epidemiological studies.

Overall, a total of 20 PRPs were evaluated:

- Twelve one-site PRPs: Six random half-mouth PRPs and six full-mouth PRPs. The single sites were MB, B, DB, ML, L, and DL.
- Two two-site PRPs: One random half-mouth PRP (as in NHANES III) and one full-mouth PRP. The two sites were MB–B.
- Four three-site PRPs: Two random half-mouth PRPs and two full-mouth PRPs. The two combinations of three sites were MB–B–DB (as in NHANES IV) and MB–B–DL.
- One random half-mouth six-site PRP. The six sites were MB–D–DB–ML–L–DL.
- One six-site PRP using the Ramfjörð teeth. The six sites were MB–D–DB–ML–L–DL.

Because they have been used in the past or because they may have particular appeal for future surveys, the authors focused on eight PRPs:

Half-mouth PRPs:

- (a) MB–B measurements (also labelled as NHANES III)
- (b) MB–B–DB measurements (also labelled as NHANES IV)
- (c) MB–B–DL measurements
- (d) MB–B–DB–ML–L–DL measurements (also termed six-site PRPs).

Full-mouth PRPs:

- (e) MB–B measurements
- (f) MB–B–DB measurements
- (g) MB–B–DL measurements
- (h) MB–B–DB–ML–L–DL measurements on six ‘Ramfjörð’ teeth—right maxillary first molar, left maxillary central incisor, left maxillary first premolar, left mandibular first molar, right mandibular central incisor, and right mandibular first premolar. MPPD, recession, and CAL were computed using all six sites per tooth for the six Ramfjörð teeth, based on a maximum of 36 possible sites per mouth. No replacements for missing Ramfjörð teeth were made.

A summary for the MCAL estimates for each PRP is presented in **Table 19.2**. The true full-mouth MCAL was 1.56 mm for this study population. The MCAL results for the four half-mouth PRPs and four full-mouth PRPs are highlighted. Biases for the multi-site PRPs MCAL estimates are all <0.1 mm, and the associated relative biases range between –4.6% and 0.9%. MCAL biases (relative biases) for the NHANES III and NHANES IV half-mouth PRP were –0.04 mm (–2.3%) and –0.05 mm (–3.4%); their corresponding full-mouth versions were –0.05 mm (–3.5%) and –0.07 mm (–4.6%). The smallest bias observed 0.01 mm (0.4%) was for the half-mouth MB–B–DL PRP estimate, and the bias for its corresponding full-mouth estimate was –0.01 mm (–0.8%). The MCAL estimate based on the Ramfjörð PRP has a 0.04-mm (2.8%) positive bias. The biases for the single-site PRPs MCAL estimates are also presented in **Table 19.2** to facilitate the interpretation of the results. MCAL estimates are larger for the lingual sites than for buccal sites. The biases (relative biases) for the single-site PRPs estimates were more variable, averaging between –0.17 mm (–11.0%) and 0.14 mm (9.2%). There are no statistically significant biases for the eight primary PRPs investigated in estimating MCAL. Any PRP whose relative bias exceeded 6% for estimating MCAL is statistically significant. Because the three sites per tooth (MB, B, DL) random half-mouth method had a very small bias in estimating disease severity and also has demonstrated high sensitivity for estimating disease prevalence (previous report for this population), it was recommended as an excellent choice for large-scale epidemiologic studies (Kingman et al. 2008).

The bias in the assessment of attachment loss is influenced by the partial recording design and the type and number of sites assessed, and is also influenced by the severity of attachment loss in the study population. These factors should be considered when selecting a partial recording method in large surveys. Furthermore, **inflation factors** (true prevalence according to full-mouth protocol/prevalence in the tested protocol) designed to adjust for the underestimation of prevalence measurements so that comparisons of

Table 19.2 Bias and relative bias for attachment loss (Kingman et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Outcome	N	Mean	SD	Bias ^a	Relative bias (%)	P-value
Half-mouth PRPs						
MB	1,437	1.41	1.77	-0.15	-9.6	0.001
B	1,437	1.64	1.73	0.08	4.9	0.097
DB	1,437	1.47	1.72	-0.09	-5.5	0.061
ML	1,437	1.57	1.95	0.01	0.7	0.833
L	1,437	1.70	1.95	0.14	9.2	0.005
DL	1,437	1.65	2.00	0.09	5.8	0.086
NHANES III	1,437	1.52	1.72	-0.04	-2.3	0.422
NHANES IV	1,437	1.51	1.71	-0.05	-3.4	0.242
MB-B-DL	1,437	1.56	1.77	0.01	0.4	0.905
6 Sites	1,437	1.57	1.79	0.01	0.9	0.766
Full-mouth PRPs						
MB	1,437	1.39	1.68	-0.17	-11.0	<0.001
B	1,437	1.62	1.66	0.06	4.0	0.149
DB	1,437	1.45	1.64	-0.11	-6.8	0.014
ML	1,437	1.57	1.88	0.01	0.4	0.896
L	1,437	1.70	1.91	0.14	8.9	0.006
DL	1,437	1.63	1.92	0.07	4.5	0.167
NHANES III	1,437	1.51	1.65	-0.05	-3.5	0.217
NHANES IV	1,437	1.49	1.64	-0.07	-4.6	0.099
MB-B-DL	1,437	1.55	1.71	-0.01	-0.8	0.779
Ramfjörð	1,430	1.60	1.81	0.04	2.8	0.358

SD standard deviation, MB mesiobuccal, B midbuccal, DB distobuccal, DL distolingual, L midlingual, ML mesiolingual, PRP partial recording protocol

^aThese are differences from the 1.56 mm true full-mouth MCAL value

results with other surveys are more meaningful (Kingman and Albandar 2002; Susin et al. 2005).

To look into this issue, let us focus on the most recent periodontal surveys. In the NHANES III (1988–1994) survey, two random quadrants and two sites per tooth (mesio-buccal and mid-buccal) were examined. If the amount of underestimation caused by the use of the particular partial examination protocol is known, an inflation factor could then be used to adjust for the bias in disease estimates so that comparisons of results with other surveys are more meaningful. To estimate the prevalence of severe AL of ≥ 5 , ≥ 6 , and ≥ 7 mm in NHANES III the observed estimates should be adjusted by a factor of 1.49, 1.75, and 2, respectively. Similarly, the corresponding findings for the NHANES 2001–2004 (recently reported by Eke et al. 2010) shows the inflation factors calculated for a partial-mouth protocol that uses two random quadrants, and three sites

per tooth are 1.12 and 1.39 to estimate the prevalence of severe AL of ≥ 3 and ≥ 6 mm, respectively. Note that the inflation factor is correlated with the severity of AL (Albandar 2011). The study of Susin et al. (2005) showed that a smaller bias should occur when using a partial-mouth protocol that uses three sites per tooth (mesio-buccal, mid-buccal, and disto-lingual) than the protocol used in the NHANES 1988–1994 that uses two random quadrants and two sites per tooth (mesio-buccal and mid-buccal). Hence, for a prevalence of AL ≥ 6 mm, the recommended inflation factor based on this protocol is 1.39 (Albandar 2011).

In periodontal surveys, the bias in disease estimates caused by partial-mouth protocols can be eliminated if a full-mouth examination is used. However, this may not be feasible because a full-mouth examination requires significantly more resources, takes longer time to achieve, and may

lead to a higher measurement error. Nevertheless, there are other ways to reduce the bias caused by partial-mouth protocols. An acceptable and feasible alternative method is to perform a full-mouth examination on a subsample of the subjects (e.g., 5% of the sample) randomly selected and a partial-mouth examination on the remaining 95% subjects. In this way, the amount of bias caused by the partial-mouth protocol can be assessed, and a more accurate and valid inflation factor can be calculated and used to correct the underestimation in periodontal disease estimates (Albandar 2011).

In addition to the above, there is also the problem of ‘clustering of data’ as several scores are collected in one subject, which has important analytical consequences. Within an individual, teeth are not independent from each other. While currently available tests (e.g., Wald test) can be applied for testing inter-subject differences (e.g., differences between women and men), standard statistical software cannot be used when intra-subject comparisons (e.g., comparisons between contralateral and antagonistic teeth) are envisaged. Statistical models applicable for dependent data are then needed. If the dependence of clustered observations is overlooked, point estimates may be similar, but variance estimates may be drastically different. Hence, the confidence intervals for the naïve model are narrower, resulting in an increased risk for type I error and erroneous rejection of the null hypothesis. Also, the multi-level nature of data (several sites per tooth and several teeth in one mouth) should not be overlooked (Leroy et al. 2010).

19.4 Definitions of Periodontitis and Methods That Have Been Used to Identify This Disease in Epidemiological Studies

Defining exactly what constitutes periodontal disease for the purpose of periodontal research is a contentious problem and is, of course, quite distinct from the diagnosis of periodontitis during clinical management of patients, for which classification criteria have previously been published (Armitage 1999; Preshaw 2009).

19.4.1 Case Definition of Periodontitis That Have Been Used to Identify This Disease

The lack of agreement on a valid case definition of periodontitis in epidemiologic studies is an important obstacle that also contributes to the uncertainty about the true prevalence of periodontitis in the United States population and hampers studies of the strength of the associations between periodontitis and systemic diseases. There are a number of reasons why consensus on a case definition of periodontitis is not easy to reach. Inconsistent examination protocols have been used in different studies (surveys, case-control studies, and so forth) for various reasons, such as the availability of limited resources. In addition, the measurement method (clinical measurements, radiographs, and so forth) and the type of variables assessed (periodontal attachment loss, probing depth, bleeding on probing, and so forth) often are different in various studies. Some studies assess local inflammation or the disease activity status. Whereas periodontal attachment loss and alveolar bone loss are valid measures of the amount of periodontal tissue loss caused by periodontitis, these two variables are not adequate measures of disease activity. Besides, although periodontitis is the main cause of periodontal attachment loss, tissue loss is also caused by other factors, such as traumatic brushing. Therefore, epidemiologic studies that report the prevalence and severity of periodontal attachment loss are indeed measuring the historical aspect of periodontitis and other causes (Albandar 2011).

Researchers have historically used an array of clinical signs and symptoms such as gingivitis, bleeding on probing, pocket depth, clinical attachment loss, radiographically assessed alveolar bone loss, and even tooth loss, the ultimate endpoint of periodontal disease. Further complications are posed by the fact that in some studies, multiple disease indicators such as pocket probing depth (PD) and clinical attachment level, both representing current pathology and cumulative tissue destruction, respectively, are used. The situation is further confused by the variation in threshold values used in defining cases regardless

of the criteria used (Leroy et al. 2010). Finally, studies have also used tooth loss—the ultimate endpoint of periodontitis—as an additional outcome variable in the context of risk assessment (Borrell and Papapanou 2005).

Recently, Savage et al. (2009) performed a systematic review and critical analysis of the definitions of periodontitis and the methods which have been used to identify and measure this disease. The methodological quality of the reviewed studies was assessed by examining criteria in relation to the studies and grouping them into generic factors and factors specific to periodontitis, respectively. The main criteria that comprised the two groups were as follows: (1) generic factors, (a) random selection/representative sample and (b) examiner calibration or training and (2) factors specific to periodontitis, (a) definition of periodontitis used, (b) Measure and/or index used, and (c) Measurement tools/probe type. These specific criteria were selected to assess the potential for bias in the reviewed studies particularly selection bias in the case of generic factors and measurement bias for factors that were specific to periodontitis.

The survey revealed heterogeneity between the studies in the measurement tools used, particularly the types of probes and the sites and areas of the mouth that were assessed (Table 19.3). There was also heterogeneity in the use of clinical attachment loss (CAL) and pocket probing depth (PPD) as criteria for periodontitis (Table 19.4). In the 15 studies, the threshold for a diagnosis of periodontitis when CAL was the criterion ranged from 2 to ≥ 6 mm and when PPD was used, from 3 to ≥ 6 mm (Table 19.5) (Savage et al. 2009).

19.4.2 Challenges Posed by the Use of Different Case Definitions Across Epidemiological Studies

As it was showed, definitions of periodontal disease vary widely in the literature, and this renders comparisons between studies impossible and greatly compromises our ability to draw meaningful conclusions from a body of published research. The corollary of this is that meta-analyses of periodontal research studies are compromised as

the number of studies that can be included in such analyses is restricted by the different inclusion criteria (and study designs) employed in different research projects (Preshaw 2009). The impact of different periodontal disease definition in epidemiological research was recently revealed by several authors.

Borrell and Papapanou (2005) reviewed the literature related to the analytical epidemiology of periodontitis generated over the past decade, focusing exclusively on analytical epidemiology issues, including the challenges posed by the use of different case definitions across studies. As is evident by the studies presented in Tables 19.6 and 19.7, these inconsistencies in the case definition affect both the current and progressive disease distribution estimates in the population and the estimates of odds ratio and relative risk and consequently, render the comparison between study findings a true challenge. Consequently, although several studies are focused on the role of the same risk factors, a direct comparison of odds ratio or relative risk between studies is hard. Another issue that needs to be accounted for in terms of case definitions is the use of full- or partial-mouth recording of parameters such as clinical attachment loss and/or pocket depth. National, large-scale epidemiologic studies have usually used partial-mouth recording methodologies, such as the system used in National Health and Nutrition Examination Survey (NHANES) III study, which examined only two sites (mesio-buccal and mid-facial) in two randomly selected quadrants under the assumption that these measurements are representative of the full-mouth status. In contrast, smaller scale studies are more likely to have used full-mouth examination methodologies (Borrell and Papapanou 2005).

Manau et al. (2008) explored whether the application of different definition criteria of periodontitis, used in other similar studies, has an influence on the significance of the association between periodontitis and prematurity or low birth weight. Fourteen periodontitis definitions and more than 50 periodontal disease continuous measurements, found in 23 published studies, were applied to a cohort study that included 1,296 pregnant women. Six of the 14 tested definitions of periodontitis resulted in statistically significant

Table 19.3 Examination methods used in the studies (Savage et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Reference	Examination area	Measurement tool	Probing location (MB, B, DB, ML, L and DL)	PPD	CAL
Borrell et al. (2005)	Half mouth	Probe (type unclear)	Mid and MB probing of randomly assigned quadrants	R	R
Laurell et al. (2003)	Full mouth	Probe (type unclear) Full-mouth radiographs measuring inter-proximal bone height	Four-point probing (location NR)	R	NR
Craig et al. (2001)	Full-mouth (third molars excluded)	North Carolina periodontal probe	Six-point probing MB, B DB, DL, L and ML	R	R
Paidi et al. (1999)	Half mouth	Williams marked probe	Six-point probing MB, B DB, DL, L and ML of randomly assigned contra-lateral diagonal quadrants	R	R
Norderyd and Hugoson (1998)	Full mouth (radiographic)	Six Bitewings and one DPT (20–30 years olds) Full-mouth intra-oral radiographs (40–80-year olds)	Location NR	NR	NR
Timmerman et al. (1998)	Part mouth	Probing depth – force-controlled probe (Borodontic Ash/Dentsply 240 N/cm ²) Attachment loss – Hu-Friedy® probe (Williams calibration)	MB, DB surfaces of all teeth except molars, as well as on vestibular and lingual surfaces of the Ramfjord teeth (16, 21, 24, 36, 41 and 44) Specific location NR	R	R
Chiappe et al. (1997)	Selected index teeth	Conventional probe (CP12 probe)	Four-point probing MB, DB, DL and ML	R	R
Anagnou-Vareltzides et al. (1996)	Full mouth	Calibrated probe with tip diameter of 0.45 mm	Six-point probing MB, B DB, DL, L and ML	R	R
Agerholm and Ashley (1996)	Full-mouth (third molars excluded)	PQW periodontal probe with Williams's markings (Hu-Friedy®)	Four-point probing MB, DB, DL and ML	R	R
Holmgren et al. (1994)	Full mouth	Community periodontal index – C periodontal probe using light probing force consistent with CPI probe	Six-point probing MB, B DB, DL, L and ML	R	R
Querna et al. (1994)	Selected index teeth	Glickman 26-G periodontal probe with round ended 0.5-mm diameter tip	Six-point probing MB, B DB, DL, L and ML	R	NR
Machtei et al. (1992)	Full-mouth (third molars excluded)	Computerized probe with standardized 20 g force	Six-point probing MB, B DB, DL, L and ML	R	R
Peng et al. (1990)	Full mouth	World Health Organization probe	Location NR	R	NR
Brown et al. (1989)	Selected index teeth	Hu-Friedy® CP6 round probe with 0.48 mm diameter	Location NR	R	NR
Wang et al. (1987)	Selected index teeth	Glickman periodontal probe	Probing on Ramfjord teeth (16, 11, 24, 36, 41 and 44) Specific location NR	R	NR

PPD pocket probing depth, MB mesiobuccal, L mid lingual, CAL clinical attachment loss, B mid buccal, DL distolingual, R recorded, DB distobuccal, NR not recorded, ML mesiolingual

adjusted odds ratios (ORs) for some of the adverse pregnancy outcomes (Table 19.8), while no significance was found for the other eight case

definitions. Out of more than 50 periodontal continuous measurements tested, only 17 demonstrated statistically significant ORs.

Table 19.4 Sample selection, periodontitis definition and tool used in the studies (Savage et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

Reference	Sample selection	Examiner calibration	Definition of periodontitis and/or index
Borrell et al. (2005)	Random and representative of general population	U	Threshold – a person who had at least three sites with clinical attachment loss ≥ 4 mm and at least two sites with PPD ≥ 3 mm. However, these conditions did not have to be present in the same site or tooth
Laurell et al. (2003)	Random and representative of general population	U	Threshold – bone loss was set at 10% of the tooth length, which corresponded to a bone loss of at least 2–3 mm
Craig et al. (2001)	Does not appear to be random and representative of general population	R	Threshold – a periodontal diseased subject was defined as having at least 20 teeth and at least four sites with pocket depths >3 mm and at least four sites with attachment loss >3 mm (based on radiographic inter-proximal loss)
Paidi et al. (1999)	Does not appear to be random and representative of general population	R	Threshold – use of three terms. (a) Prevalence of LOA: as the percentage of LOA of x mm (where $x=2, 4, 6$ and 9 mm). (b) Extent of LOA: as the mean percentage of sites with LOA of x mm or more per person. (c) Severity: LOA of x mm or more per person. Threshold defining periodontitis – NR
Norderyd and Hugoson (1998)	Does not appear to be random and representative of general population	R	Threshold – criteria from Hugoson and Jordan (1982). Group 3 – alveolar bone loss around the majority of the teeth not exceeding 1/3 of the length of roots, Group 4 – alveolar bone loss around the majority of the teeth ranging between 1/3 and 2/3 of the length of the roots. Group 5 – alveolar bone loss around the majority of the teeth exceeding 2/3 of the length of the roots; presence of angular bony defects and/or furcation defects
Timmerman et al. (1998)	Representative study of community	R	Threshold – criteria from Brown et al. (1990). No or minor periodontitis – 0–2 mm maximum attachment loss, moderate periodontitis – 3–4 mm maximum attachment loss, advanced periodontitis ≥ 5 mm maximum attachment loss
Chiappe et al. (1997)	Does not appear to be random and representative of general population	R	Threshold – loss of attachment was determined when the clinical attachment loss was ≥ 2 mm
Anagnou-Vareltzides et al. (1996)	Random and representative of general population	U	Threshold – level of ≥ 6 mm for pocket probing depth and periodontal attachment level was utilized as expressing deep pocketing and advanced attachment losses as an arbitrary definition of severe periodontal disease
Agerholm and Ashley (1996)	Does not appear to be random and representative of general population	U	Threshold – clinical attachment loss at 2, 3 or 4 mm. Subsets made up of ten index teeth recommended by the WHO for partial recording (two molars in each quadrant and maxillary right and mandibular left central incisors) and a subset comprising maxillary buccal and mandibular lingual sites ('Pritchard sites')
Holmgren et al. (1994)	Does not appear to be random and representative of general population	R	Threshold – definition based on ≥ 6 and 9 mm loss of attachment
Querna et al. (1994)	Does not appear to be random and representative of general population	NR	Threshold – subjects with inflammation and PPD of over 3 mm but <5 mm were categorized as early periodontitis Over 5 mm moderate to advanced periodontitis periodontal screening exam (PSE) index
Machtei et al. (1992)	Does not appear to be random and representative of general population. 'Convenience' sample also used	R	Threshold – the clinical entity of established periodontitis is suggested based on the presence of clinical attachment level ≥ 6 mm in two or more teeth and one or more sites with Pocket probing depth ≥ 5 mm

Table 19.4 (continued)

Reference	Sample selection	Examiner calibration	Definition of periodontitis and/or index
Peng et al. (1990)	Random and representative of general population	R	Threshold – periodontitis included teeth with pocketing and attachment loss – CPITN 3 – pocket depth between 3.5 and 5.5 mm, CPITN 4 – pocket depth >5.5 mm Definition made according to the modified method of Gaengler (1984)
Brown et al. (1989)	Random and representative of general population	R	Threshold – periodontitis (pockets ≥ 4 mm) Advanced periodontitis (pockets ≥ 6 mm) Definition based on modified Russell's periodontal index
Wang et al. (1987)	Not random and representative of general population	NR	Threshold – periodontal disease index above three was considered to suffer from periodontitis Definition based on Ramfjord periodontal disease index 1959

R recorded, *NR* not recorded, *U* unclear, *PPD* pocket probing depth, *CAL* clinical attachment loss, *LOA* loss of attachment, *CPITN* community periodontal index of treatment need

Table 19.5 Numerical ranges of CAL and PPD thresholds used in the studies (Savage et al. 2009) (Reprinted with permission from John Wiley & Sons, Inc.)

≤ 2 mm	≥ 2 mm	≥ 3 mm	≥ 4 mm	≥ 5 mm	≥ 6 mm
<i>CAL study</i>					
Timmerman et al. (1998) (no or minor periodontitis) at one or more sites	Agerholm and Ashley (1996) at one or two approximal sites Chiappe et al. (1997) at one site	Craig et al. (2001) at four sites Timmerman et al. (1998) (moderate periodontitis) at one or more sites Agerholm and Ashley (1996) at one or two approximal sites	Borrell et al. (2005) at three sites Agerholm and Ashley (1996) at one or two approximal sites	Timmerman et al. (1998) (advanced periodontitis) at one or more sites	Anagnou-Vareltzides et al. (1996) (severe periodontitis) at mean sites Holmgren et al. (1994) at mean sites Machtei et al. (1992) at two or more teeth
<i>PPD study</i>					
		Borrell et al. (2005) at two sites Craig et al. (2001) at four sites Peng et al. (1990) (3.5 mm CPITN) worst individual score at sextant	Querna et al. (1994) (early) worst individual score at sextant. Brown et al. (1989) (modified Russell index) at one site	Querna et al. (1994) (moderate – advanced) at the highest score per sextant Machtei et al. (1992) at one or more teeth Peng et al. (1990) (5.5 mm CPITN) worst individual score at sextant	Anagnou-Vareltzides et al. (1996) (severe periodontitis) at mean sites Brown et al. (1989) advanced periodontitis (modified Russell index) at one site

CAL clinical attachment loss, *CPITN* community periodontal index of treatment need, *PPD* pocket probing depth

Similar results were reported by Kassab et al. (2011) who studied the impact of using various definitions of periodontitis on the frequency of periodontitis and on the associations with some known risk factors for periodontitis in a population of post-partum women in France. A clinical

examination was performed within 2–4 days post-partum in 932 at-term women at five maternity units. Six definitions were used for periodontitis, using single or combined criteria. Five definitions required the involvement of at least two teeth. The sixth definition required the involvement of at

Table 19.6 Analytical epidemiology of periodontitis: cross-sectional and case-control studies (Borrell and Papapanou 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Grossi et al. (1994)	To identify risk indicators and putative risk factors for periodontitis Cross-sectional study (US – Erie County, NY): $N=1,426$ adults aged 25–74 years. Mostly whites (90.3%)	Mean CAL and classification as: healthy (≤ 1 mm CAL); low (>1 and ≤ 2 mm); moderate (>2 and ≤ 3 mm); high (>3 but ≤ 4 mm); and severe (≥ 4 mm)	Periodontal status: supragingival plaque, GB, subgingival calculus, PD, CAL. Subgingival microbiota: <i>Actinobacillus actinomycetemcomitans</i> , <i>Tannerella forsythensis</i> , <i>Campylobacter rectus</i> , <i>E. saburreum</i> , <i>Fusobacterium nucleatum</i> , <i>Porphyromonas gingivalis</i> , <i>capnocytophaga</i> spp. and <i>Prevotella intermedia</i> . Other covariates: age, gender, race, education, income, smoking and numbers of packs/year, exposure to occupational hazards, systemic diseases	In a multivariable logistic model, age (ORs ranging from 1.72 for age group 35–44–9.01 for age group 65–74), male gender (OR = 1.36; 95% CI: 1.06–1.76), smoking (ORs ranging from 2.05 [95% CI: 1.47–2.87] for light smokers to 4.75 [95% CI: 3.28–6.91] for heavy smokers), diabetes (OR = 2.32; 95% CI: 1.17–4.60), <i>P. gingivalis</i> (OR = 1.59; 95% CI: 1.11–2.25) and <i>T. forsythensis</i> (OR = 2.45; 95% CI: 1.87–3.24) were positively associated with severity of CAL, while <i>capnocytophaga</i> spp. (OR = 0.60; 95% CI: 0.43–0.84), anaemia (OR = 0.65; 95% CI: 0.42–0.99), allergy (OR = 0.77; 95% CI: 0.58–1.00) and education (OR = 0.65; 95% CI: 0.51–0.85) were protective against severe CAL
Grossi et al. (1995)	To identify factors associated with alveolar BL. Cross-sectional study (US – Erie County, NY): $N=1,361$ aged 25–74 years. Mostly whites	Radiographic alveolar BL: distance from CEJ to the alveolar crest. Inter-proximal BL: healthy (0.4, <2 mm), low (2.0, <3.0 mm), moderate (3.0, <4.0 mm), and severe (≥ 4.0 mm)	Supragingival plaque, GB, subgingival calculus, PD, CAL. <i>A. actinomycetemcomitans</i> , <i>T. forsythensis</i> , <i>C. rectus</i> , <i>Eubacterium saburreum</i> , <i>F. nucleatum</i> , <i>P. gingivalis</i> , <i>capnocytophaga</i> spp., and <i>P. intermedia</i> . Other covariates: age, gender, race, education, income, smoking, and packyears, exposure to occupational hazards, systemic diseases	In a multivariable logistic regression model, age (ORs ranging from 2.6 [95% CI: 1.75–3.83] for age group 35–44 to 24.08 [95% CI: 15.93–36.29] for age group 65–74), male gender (OR = 1.29; 95% CI: 1.05–1.61), race other than white (OR = 2.40; 95% CI: 1.21–4.79), smoking (ORs ranging from 1.48 [95% CI: 1.02–2.14] for light smokers to 7.28 [95% CI: 5.09–10.31] for heavy smokers), <i>P. gingivalis</i> (OR = 1.73; 95% CI: 1.27–2.37), <i>T. forsythensis</i> (OR = 2.52; 95% CI: 1.98–3.17), were significantly associated with increasing severity of BL, while kidney disease (OR = 0.55; 95% CI: 0.35–0.85), allergy (OR = 0.76; 95% CI: 0.59–0.98) and education (OR = 0.67; 95% CI: 0.53–0.98), significantly decreased the severity of BL.
Martinez Canut et al. (1995)	To evaluate the influence of tobacco on the severity of periodontal disease and to quantify the strength of the influence in relation to the amount of tobacco consumed. Cross-sectional study (Spain): $N=889$ subjects with mild and advanced periodontitis, including 52.6% smokers	GR, PD, CAL, and tooth mobility	Smoking status (yes/no), and categorization based on daily cigarette consumption. Other covariates: age, sex	Daily cigarette consumption had a dose-response effect on GR, PD, CAL, and mobility. The effect of age and tobacco on mobility was comparable for adults older than 51 years of age and consuming more than 21 cigarettes/day (0.76 and 0.82, respectively). Tobacco consumption had a dose-response effect on CAL across all ages

Bridges et al. (1996)	To determine differences in plaque and gingival indices, bleeding scores, PDs, loss of attachment, and number of missing teeth between diabetic and non-diabetic subjects. Case-control study (US): $N=233$ men aged 24–78 years. One hundred and eighteen diabetic (46 type 1 and 72 type 2) and 115 non-diabetic subjects, matched for age and BMI	PI, GI, bleeding on probing, PD, CAL, and number of missing teeth	Diabetes mellitus Type 1 and Type 2. Glycosylated haemoglobin, fasting blood glucose, serum insulin, and C-peptide levels. Covariates: age, income, education, brushing, flossing, dental care, smoking status	PI ($P<0.0001$); GI ($P<0.0002$); bleeding score ($P<0.0001$); PD ($P=0.0059$); loss of attachment ($P<0.0001$); and missing teeth ($P<0.005$) were significantly higher in diabetic than non-diabetic men. Income and education were negatively associated with periodontal status (plaque, gingival indices, and missing teeth with income; and plaque and gingival indices, PD, CAL, and missing teeth for education) for both diabetic and non-diabetic men
Alpagot et al. (1996)	To evaluate risk indicators of periodontal disease. Cross-sectional study (US): $N=117$ subjects 18 years and older.	Healthy sites: GI=0, PD $\leq$ 3 mm and CAL1 mm; gingivitis sites: GI>0, PD $\leq$ 3 mm and CAL $\leq$ 2 mm; periodontitis sites: GI>0, PD>3 mm and CAL>+3 mm. Subjects were assigned to periodontitis and gingivitis groups based on the subject's most DS per sextant	GCF, plaque, β G, NE, MPO, <i>F. nucleatum</i> , <i>P. intermedia</i> , and <i>P. gingivalis</i> . Other covariates: age, race (white, black, native American, and Asian), cigarette smoking (pack/years)	<i>F. nucleatum</i> , <i>P. gingivalis</i> , and <i>P. intermedia</i> were associated with PD and CAL. Age, race, smoking pack/years, β G, NE, MPO, <i>F. nucleatum</i> , <i>P. gingivalis</i> , and <i>P. intermedia</i> were risk indicators for periodontitis in this racially diverse urban population
Kornman et al. (1997)	To investigate the association of a specific polymorphism in the IL-1 gene cluster and the severity of periodontitis. Case-control study (US): $N=134$ aged 35 and older whites only.	Periodontitis: mild ($n=49$); no PD 3+ mm and no sites with BL>15% of the root length. Moderate ($n=42$); <4 inter-proximal sites with $\geq 50\%$ BL and a total mean BL of 17–28%; Severe ($n=43$); ≥ 7 inter-proximal sites with $\geq 50\%$ BL and total mean BL of >34%.	SNPs at the following positions: IL-1A(–889); IL-1B(–511); IL-1B(+3953); IL-1RN (intron 2) and TNF-A –308. Other covariates: smoking (current or not).	Among non-smokers, there was a strong association between severity of periodontitis and the composite genotype comprising allele 2 at position IL-1A –889 plus allele 2 at position +3953 of the IL-1B gene (OR 6.8; 95% CI: 1.01–45.62). In non-smokers aged 40–60 years, the composite genotype was present in 78% of the subjects with severe ($n=9$), 26% of those with moderate ($n=30$), and 16% of those mild BL ($n=32$) (OR 18.9; 95% CI: 1.04–343.05).

(continued)

Table 19.6 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Gore et al. (1998)	To determine the distribution of specific polymorphisms in the IL-1 gene cluster in patients with adult periodontitis and their matched controls; and to examine possible association of genotype with cytokine production by oral and peripheral blood PMN. Case-control study (US): $N = 64$ (32 cases and 32 age- and sex-matched 32 controls) white	Periodontal disease: early ($n = 10$); moderate ($n = 10$); and advanced ($n = 12$)	SNPs at the following positions: IL-1B +3953, IL-1B -511, IL-1A -889, IL-1RN, assessed at oral and peripheral blood PMN. Smoking status of controls was not determined.	Homo- and heterozygosity with the rare allele 2 at position IL-1B +3953 (1/2 and 2/2) was significantly more frequent among cases with advanced adult periodontitis than among early and moderate periodontitis. Peripheral blood PMNs of these patients produced increased levels of the IL-1 β compared with patients without allele 2, although this increase was not statistically significant. The composite genotype (IL-1B +3953 allele 2 and IL-1A -889 allele 2) was carried by 50% of cases with advanced periodontitis, 25% of cases with early/moderate periodontitis, and 28% of controls. However, this difference did not reach statistical significance
Boström et al. (1998a)	To evaluate the influence of smoking on TNF- α levels in GCF of patients with untreated moderate-to-severe periodontal disease (30 current smokers, 19 former smokers, and 29 non-smokers, 31–79 years old). Cross-sectional study (Sweden): $N = 78$	TNF- α in GCF	Smoking status: current (mean 19.3 cigarettes/day for an average of 25.4 years), former (19.5 cigarettes/day for 17 years) and never. PI, GB, mean PD. Sites with PD ≥ 4 mm were defined as DSs and were expressed as percentage. Other covariates: age, gender, and presence in dental plaque of <i>A. actinomycetemcomitans</i> , <i>P. gingivalis</i> , <i>intermedia</i>	TNF- α levels were significantly higher in former and current smokers when compared with never smokers. TNF- α (log) was associated with current and former smoking status and PI after adjusting for GB and DS
Boström et al. (1999)	To examine the influence of tobacco smoking on GCF levels of IL-6, IgA, IgG, and TNF- α in patients with moderate-to-severe periodontal disease. Cross-sectional study (Sweden): $N = 108$, age range 30–81 years	IL-6, IgA, IgG, and TNF- α in GCF	Current (42%, mean 21 cigarettes/day for 24 years), former (26%, mean 20 cigarettes/day for 18 years), and never smokers (32%). PL, GB, PD. Other covariates: presence in dental plaque of <i>A. actinomycetemcomitans</i> , <i>P. gingivalis</i> , and <i>P. intermedia</i>	GCF IL-6 was not associated with smoking. Levels of TNF- α were significantly higher among smokers when compared with non-smokers. The ratio of TNF- α /albumin was statistically greater in current smokers as compared with non-smokers. IgA/IgG and IgA/albumin were found to be correlated in all subjects. There was no significant difference in the level of the investigated pathogens across smoking groups

Galbraith et al. (1999)	To examine the distribution of SNPs at the IL-1B +3953 and TNF-A -308 in adults varying levels of periodontal disease; and to examine the influence of the genotype on cytokine production by oral and peripheral blood PMNs. Cross-sectional study (US): N = 85 (20 with advanced periodontitis, 20 with plaque-associated gingivitis, and 45 referent subjects). Whites, aged 35–65 years old	Advanced periodontal disease: ≥ 3 sites with PD ≥ 7 mm and CAL > 5 mm. Oral and peripheral blood PMN IL-1 β and TNF- α	SNPs at IL-1B +3953 and TNF-A -308. Bleeding, mobility, smoking	Homo- and heterozygosity for the rare IL-1B +3953 allele 2 (1,2 and 2,2) was significantly elevated in the patient group compared with the referent group. TNF-A -308 1,1 genotype was significantly elevated in patients with advanced periodontitis compared with those with gingivitis. Allele 1 frequency was higher in patients with gingivitis and advanced periodontitis compared with the referent group. No significant differences were found in IL-1 β production by LPS-stimulated peripheral blood PMNs, or in TNF- α synthesis by either oral or peripheral blood PMNs. Increased IL-1 β production by oral PMNs was observed in carriers of allele 2 with advanced periodontal disease.
Genco et al. (1999)	To investigate the relationship of periodontal disease to stress, distress, and coping behaviours in a large population-based sample of adults. Cross-sectional study (US): N = 1,426 aged 25–74 years. Whites (95.6%)	Periodontal status was categorized by mean CAL as healthy (0–1 mm), low-level disease (1.1–2), moderate (2.1–3), high (3.3–4) and severe (4.1–8); and by mean ACH as healthy (0.4–1.9 mm), low-level disease (2–2.9), moderate (3–3.9), and severe (4).	Life Event Scale, measures of chronic stress (Problems of Everyday Living Scale), measures of distress, coping styles and strategies, and hassles and uplifts. Other covariates: age, gender, years of education, smoking, income, diabetes, allergy, anaemia. <i>A. actinomycetancomitans</i> , <i>T. forsythensis</i> , <i>C. rectus</i> , <i>Capnocytophaga</i> species, <i>E. saburreum</i> , <i>F. nucleatum</i> , <i>P. gingivalis</i> , and <i>P. intermedia</i>	Subjects with high emotion-focused coping and more financial strain were twice as likely to have more severe CAL (OR: 2.24; 95% CI: 1.15–4.38) and ACH (OR: 1.91; 95% CI: 1.15–3.17) than those with less financial strain. Similar patterns were observed for low problem-focused copers (CAL: OR = 2.21; 95% CI: 1.11–4.38; and ACH: OR = 2.12; 95% CI: 1.28–3.51). The risk for periodontal disease as measured by CAL and ACH was higher in subjects with high financial strain and high emotion-focused coping (poor coping). High financial strain accompanied by high levels of problem-focused coping (good coping) entailed no risk for periodontitis

(continued)

Table 19.6 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
McDevitt et al. (2000)	To evaluate the association between the IL-1 gene polymorphism in a broad spectrum of patients typically found in a dental practice. Case-control study (US): $N=90$ (44 cases and 46 controls) adults 35 years and older	Periodontal status: mean inter-proximal BL. Cases: BL ≥ 3.0 mm or at least 10 teeth with at least one site with ≥ 3.0 mm BL and at least two sites in each quadrant with ≥ 3.0 mm BL. Controls: no more than 2 sites with BL ≥ 3.0 mm	SNPs at positions IL-1A +4845, IL-1B +3954. Other covariates: age, gender, ancestry, smoking status	Patients with moderate-to-severe periodontal disease were more likely (41%) to be IL-1 genotype positive according to Kornman et al. (1997) compared with patients in the healthy and mild group (28%), but the difference did not reach statistical significance. Age (OR = 1.26, $P < 0.001$), smoking history (OR = 7.43, $P = 0.031$) and being IL-1 genotype positive (OR = 3.75) were significantly associated with severity of periodontal disease. There was a significant interaction between smoking history and positive IL-1 genotype (OR = 1.68, $P = 0.045$). In a separate analysis conducted for patients of European origin, the OR for those IL-1 genotype positive was stronger (OR = 5.27, $P = 0.026$) than that observed for the total population
Parkhill et al. (2000)	To explore the association of IL-1 gene polymorphisms with EOP. Case-control (UK): $N=70$ with EOP (including 21 with localized EOP) and 72 without clinical evidence of periodontitis Caucasians	EOP: clinical and radiographic evidence of rapid rate of tissue destruction with age of onset < 35 years. L-EOP: ≥ 3 mm of BL confined only to incisors and first molars and adjacent tooth surfaces. Controls: CAL ≤ 1 mm	SNPs at positions IL-1B +3953 and IL-1RA. Smoking	When compared with the control group, subjects with EOP were 2.22 more likely to carry the IL-1B +3953 I/I genotype. There was a significant difference in the IL-1B +3953 allele distribution between EOP smokers and control smokers. There was no difference in the frequency of IL-1RA allele occurrence or genotype between EOP and control smokers, or EOP and control non-smokers

Bergström et al. (2000b)	To examine the association of pocket depth and alveolar bone height with smoking. Cross-sectional study 1992–1993 (Sweden): $N=244$ aged 20–69 years	Pocket depth ≥ 4 mm defined a DS and ≥ 6 mm a severely DS. GB and PI. Periodontal bone height was calculated as a mean across all measured inter-proximal sites	Smoking exposure was defined in terms of consumption (number of cigarettes/day), duration (number of years of smoking), and lifetime exposure (Product of daily consumption, and years of duration-cigarette/years). Heavy versus light consumption: ≥ 10 cigarettes/day versus <10 cigarettes/day Duration: ≥ 15 years versus <15 years Life-time exposure: ≥ 200 cigarette/years versus <200 cigarette/years. Other covariates: age	Compared with former and non-smokers, current smokers had the highest prevalence of DSs (≥ 4 mm) with older (40–69) current smokers (27.0%) having a significantly higher prevalence than younger (20–39) current smokers (3.8%). This pattern was observed when comparing heavy versus light smokers according to consumption, duration, and lifetime exposure. In terms of periodontal bone height, there was no difference among current, former and non-smokers in the younger group (20–39). However, in the older group (40–69), current smokers had lower bone height than non-smokers. In multiple regression, life-time exposure ($P<0.001$) was associated with the frequency of DSs and periodontal bone height after adjusting for age, GB, and PI
Armitage et al. (2000)	To examine the association between the composite genotype of IL-1A +4845 and IL-1A +3954 (formerly +3953) and the severity of periodontal disease in a population of Chinese heritage. Cross-sectional study (US): $N=300$, 21–69 years old	PI, bleeding on probing, PD, CAL and GR. Healthy or gingivitis: mean $CAL \leq 0.5$ and no inter-proximal sites with $CAL \geq 3$ mm. Initial: mean $CAL \geq 0.6$ –1.5 mm, no inter-proximal sites with $CAL \geq 3$ mm; Moderate: mean $CAL = 1.6$ –2.4 mm and ≤ 8 sites with inter-proximal $CAL \geq 3$ mm distributed through at least three quadrants or at least 6 teeth; Severe: mean $CAL \geq 2.5$ and ≥ 1 site with $CAL \geq 5$ mm in three out of four quadrants	SNPs at positions IL-1A +4845 and IL-1A +3954. Other covariates: age, gender, ethnicity, smoking status: current, former and never cigarette consumption: pack/day $\times$ number of years	The composite genotype according to Parkhill et al. (2000) was present in 2.3% (7/300) of the subjects. Allele 2 at position IL-1A +4845 was carried by 17% of the population, of whom only two were homozygous. Allele 2 of IL-1B +3954 was carried by only 3.3% of the study population and all subjects were heterozygous

(continued)

Table 19.6 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Boström et al. (2000)	To examine the influence of smoking on the GCF levels of IL-1 β and IL-1ra in a population of patients with established periodontitis. Levels of IgA and IgG in GCF were also determined. Cross-sectional study (Sweden): N=40 aged 32–86 years	PL, GB, PPdall, DS, PPdDs, DSs: sites with PD $\geq$ 4 mm PPdall: mean across all sites with PD. PPdDs: mean PD of DSs	Smoking status: current (20 cigarettes/day for and average of 31 years), former, and non-smokers. Other covariates: age and gender. GCF: IL-1 β , IL-1ra, IgA, IgG, albumin, protein	Neither IL-1 β nor IL-1ra was significantly different in smokers and non-smokers. The same was the case for IgA, IgG, albumin, and protein. IL-1 β and IL-1ra were correlated among smokers. IL-1 β was correlated with albumin and protein in smokers and non-smokers. Correlations among smokers and non-smokers were found between IgA and IgG; IgA and albumin; IgG and albumin; IgA and protein and IgG and protein. However, these correlations were not different between smokers and non-smokers.
Albandar et al. (2000)	To examine whether the association between cigar and pipe smoking and periodontal disease is similar to the one observed with cigarette smoking; and whether cigar, pipe, and cigarette smoking are associated with tooth loss. Cross-sectional study (USA): N=705 aged 21–92 years, 52% males and 87% white	Subjects with $\geq$ 8 teeth or $\geq$ 50% of teeth with GB were classified as having extensive gingival inflammation; those with 3–7 teeth or 25–49% of their teeth as with limited gingival inflammation. Periodontitis: advanced: PD $\geq$ 5 mm in 4+ teeth (or $\geq$ 30% of teeth), or PD $\geq$ 4 mm in 8+ teeth (or $\geq$ 60% of teeth); moderate: PD $\geq$ 5 mm in 2–3 teeth or PD $\geq$ 4 mm in 4–7 teeth (or 30–59% of teeth); mild: PD $\geq$ 4 mm in 1+ teeth or PD $\geq$ 3 mm in 2+ teeth	Cigar, pipe and cigarette smoking: current, former, and never. Other covariates: age, gender, and race	Current cigarette smokers exhibited the worst periodontal status followed by former smokers. Current (25.7%) and former (20.2%) cigarette smokers had higher prevalence of moderate/advanced periodontitis than non-smokers (13.1%). However, current cigarette smokers had lower GB than non-smokers (6.6% versus 17.9%, $P<0.08$). Former cigarette smokers (13.6%) also had lower GB than non-smokers. Calculus was more prevalent among current cigarette smokers. Current or former pipe or cigar smokers had poorer periodontal conditions than non-smokers as indicated by higher prevalence of moderate/advanced periodontitis (17.6% versus 6.1%, $P<0.006$). As with cigarette smoking, current or former pipe and cigar smokers had lower prevalence of GB (8% versus 15.7%) but lower subgingival calculus. In multiple linear regression, current and former smoking, regardless of type, was associated with increased percentage of subjects with moderate/advanced periodontitis after adjusting for age, gender, race, and numbers of years of being smoking cigarette, cigar and pipe. Current smoking was also associated with higher number of missing teeth

Boström et al. (2001)	To describe the detection rates of selected bacterial species and analyse their inter-relationships in smokers and non-smokers. Case-control study convenience sample (Sweden): $N=64$ aged 36–86 years. 33 smokers and 31 non-smokers	Subgingival <i>P. gingivalis</i> , <i>P. intermedia</i> , <i>Prevotella nigrescens</i> , <i>T. forsythensis</i> , <i>A. actinomycetemcomitans</i> , <i>F. nucleatum</i> , <i>Treponella denticola</i> , <i>Peptostreptococcus micros</i> , <i>C. rectus</i> , <i>Eikenella corrodens</i> , <i>Selenomonas noxia</i> and <i>Streptococcus intermedius</i> . Checkerboard hybridizations non-colonized versus colonized. Non-colonized and less heavily colonized versus heavily	PD (mean score for four sites: buccal, distal, lingual, and mesial). GB (yes/no within 30 sites in response to probing with a pressure of approximately 0.25 N). Other covariates: age and gender	Using score 1 as cutoff for presence, contrasting colonized versus non-colonized patients, 8 out of 12 species were detected in $\geq 90\%$ of both smokers and non-smokers. Using score 4 as cut-off, contrasting heavily colonized patients versus non-colonized and less heavily colonized patients, detection rates decreased in both smokers and non-smokers. No significant differences in detection rates were observed between smokers and non-smokers. Logistic regression analysis indicated that neither smoking, PD nor GB influenced the occurrence of the species analysed
Laine et al. (2001)	To investigate the distribution of polymorphisms in the genes of the IL-1 cluster among periodontitis patients and controls after adjusting for selected risk factors. Case-control study Convenience sample (the Netherlands): $N=105$ (53 cases and 53 controls) aged 25 years and older. Caucasians	Cases: severe periodontitis ≥ 7 inter-proximal sites with 50% BL. Controls: subjects with <4 mm PD and no sites with BL	SNPs at IL-1A +3954, IL-1B and IL-1RN. Smoking (current and never). <i>A. actinomycetemcomitans</i> and <i>P. gingivalis</i>	There were no significant differences in the carriage of the alleles 1 and 2 between cases and controls regardless of their smoking status. When smokers and non-smokers were divided based on the presence or absence of <i>A. actinomycetemcomitans</i> and <i>P. gingivalis</i> , non-smoker cases who were negative for <i>P. gingivalis</i> were found to carry significantly more often the combination of alleles IL-1A #2 and IL-1RN #2 than controls (32.1% versus 11.3%, $P=0.034$). This was also true for <i>A. actinomycetemcomitans</i> negative cases (32.5% versus 11.3%, $P=0.018$). Furthermore, subjects were divided in negative and positive for both <i>A. actinomycetemcomitans</i> and <i>P. gingivalis</i> . Smoker patients negative for both pathogens were more likely to carry alleles IL-1A #2, IL-1B #2 and IL-1RN #2 than controls (42.1% versus 11.3%, $P=0.0068$).

(continued)

Table 19.6 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Papapanou et al. (2001)	To examine the association of the IL-1 gene polymorphism and certain elements of periodontal status. Case-control study (Sweden): $N=205$ aged >20 years 132 cases and 73 controls age- and gender-matched	Cases: moderate-to-advanced periodontal disease, with several inflamed periodontal pockets, loss of attachment, and radiographic evidence of BL. Controls: no or only minimal CAL despite plaque accumulation and bleeding on probing. Presence of 19 bacterial species subgingivally and levels of serum IgG to these bacteria	SNPs at IL-1A (+4845) and IL-1B (+3954). Age, gender, smoking	The positive polymorphism was more prevalent in cases (26.8%) compared with controls (16.1%) but the difference was not statistically significant ($P=0.624$). The positive polymorphism was associated with an increase in percentage of sites with AL ≥ 6 mm (modelled as continuous and discrete) and an overall mean serum antibody response (discrete) in cases. Among non-smokers, the presence of the positive polymorphism was associated with an increase percentage of sites with AL ≥ 6 mm and mean CAL
Papapanou et al. (2002)	(1) To examine the association of subgingival bacterial profiles to clinical periodontal status in a population with limited access to dental care; (2) to examine the external validity of the findings from an earlier study (Papapanou et al. 1997). Cross-sectional study (Thailand): $N=356$, 30-39 and 50-59 years old	PD and CAL were assessed at 6 sites/tooth, at all teeth apart from third molars. Subjects were grouped according to different levels of pocketing/attachment loss: subjects with ≥ 3 sites with PD ≥ 5 mm (59%, G1); ≥ 10 sites with CAL ≥ 5 mm (50%, G2); and ≥ 30 sites with CAL ≥ 5 mm (24%, G3)	Subgingival plaque samples were obtained at maximally 14 sites/subject. Checkerboard hybridizations were used to analyse a total of 4343 samples with respect to 27 bacterial species	ORs for heavy colonization (i.e., bacterial levels exceeding the specific thresholds generated in the earlier study by Papapanou et al. (1997) by 'red complex' species (<i>P. gingivalis</i> , <i>T. forsythensis</i> , <i>T. denticola</i>) were 3.7 (95% CI: 2.3-5.9) for G1; 4.0 (95% CI: 2.5-6.6) for G2; and 4.3 (95% CI: 2.6-7.1) for G3. ORs for heavy colonization by selected 'orange complex' species (<i>F. nucleatum</i> , <i>P. intermedia</i> , <i>P. nigrescens</i> , <i>P. micros</i> , <i>E. nodatum</i> , <i>C. rectus</i> , and <i>Campylobacter showae</i>) were 1.5 (95% CI: 0.8-2.9) for G1; 1.5 (95% CI: 0.8-2.9) for G2; and 1.5 (95% CI: 0.8-3.1) for G
Aleksejunene et al. (2002)	To assess the relationship between psychosocial stress and lifestyle and periodontal status in Lithuanian adults. Cross-sectional study (Lithuania): $N=681$ aged 35-74	Periodontal disease was assessed by means of the Community Periodontal Index of Treatment Needs (CPITN), including measures of PD and recession	Social support, stress, nervousness, coping, smoking, oral hygiene. Other covariates: age and gender	No association between psychosocial stress and periodontitis. However, lifestyle factors were associated with periodontitis

Meisel et al. (2003)	To examine the role of IL-1 polymorphism as a risk factor for periodontal diseases in a randomly selected population. Interaction between smoking and presence of genotype was also examined. Cross-sectional (Germany); $N = 1,085$	Percentage of sites with CAL ≥ 4 mm: subjects at the upper quintiles were compared with those at the lower quintiles	SNPs at IL-1A (-889) and IL-1B (+3954) Subjects heterozygous and homozygous for the #2 allele at each of the 2 sites analysed were designated as genotype positive. Age, gender, plaque, education, smoking, age at which the subject start smoking, frequency and duration of smoking	Periodontal disease was associated with smokers with positive IL-1 genotype after adjusting for age, sex, education, and plaque (OR = 2.5; 95% CI: 1.21–5.13). There was no association among non-smokers. Smokers had lost more teeth than their non-smoker counterparts regardless of their IL-1 genotype
Al-Zahrani et al. (2003)	To examine the association between obesity and periodontal disease. Cross-sectional study (US); $N = 13,665$ 18+ years of age. Non-Hispanic blacks and whites, Mexican Americans, and others	Periodontitis defined as at least 1 site with CAL ≥ 3 and PD ≥ 4	Obesity: BMI and WC. Other covariates: age, race, gender, poverty index, education, time since last dental visit, smoking and diabetes	A significant association between measures of body fat and periodontitis was found among the younger adults, but not middle-aged or older adults. Adjusted ORs for having periodontitis were 0.21 (0.080–0.565), 1.00 (0.705–1.407), and 1.76 (1.187–2.612) for subjects with BMI < 18.5, 25–29.9, and ≥ 30 kg/m ² , respectively. Young subjects with high WC had an adjusted OR of 2.27 (1.480–3.487) for periodontitis
Susin et al. (2004)	To estimate the number and percentage of cases with severe attachment loss attributable to cigarette smoking in a representative adult urban population in southern Brazil. Cross-sectional study (Brazil); $N = 974$ aged 30–103 years	Periodontal status: presence of severe attachment loss, defined as subjects with CAL ≥ 5 mm in $\geq 30\%$ of the teeth	Smoking status: current and former lifetime consumption (# of cigarettes/day $\times$ # of days of habit/20). Smoker: heavy ($>7,300$ packs); moderate (2,735–7,300), light (1–2,374), and non-smokers (<1 pack). Other covariates: race (white or non-white), SES, calculus ($<25\%$, 25–50%, and $>50\%$)	Nearly half of the population had CAL ≥ 5 mm at $\geq 30\%$ of their teeth (49.7%) and had been exposed to cigarette smoking (50.9%). Heavy and moderate smokers had significantly higher prevalence of CAL ≥ 5 mm than non-smokers. This finding was persistent in multivariate analysis with heavy (OR = 3.6; 95% CI: 2.2–6.0) and moderate smokers (OR = 2.0; 95% CI: 1.4–2.9) having higher odds for CAL than non-smokers. The attributable fraction of CAL because of cigarette smoking was 37.7% and 15.6% among heavy and moderate smokers, respectively.

PD probing depth, *CAL* clinical attachment level, *OR* odds ratio, *CI* confidence interval, *BL* bone loss, *CEJ* cemento-enamel junction, *GR* gingival recession, *BMI* body mass index, *GI* gingival index, *GCF* gingival crevicular fluid, *bG* β -glucuronidase, *NE* neutrophil elastase, *MPO* myeloperoxidase, *IL* interleukin, *SNP* single nucleotide polymorphism, *TNF* tumour necrosis factor, *PMN* polymorphonuclear cells, *PI* plaque index, *GB* gingival bleeding, *DS* diseased site, *IgG* immunoglobulin G, *LPS* lipopolysaccharide, *ACH* alveolar crestal bone height, *EOP* early onset periodontitis, *L-EOP* localized EOP, *CPITN* Community Periodontal Index of Treatment Needs, *WC* waist circumference, *NHANES* National Health and Nutrition Examination Survey

Table 19.7 Analytical epidemiology of periodontitis: Longitudinal studies (Borrell and Papapanou 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Beck et al. (1995)	To present the incidence of CAL in subjects who show progressive CAL in sites previously without disease and subjects who experience further progression at sites with pre-existing disease; and to compare and contrast the characteristics of people with the two types of progressive CAL. Exams at 18 and 36 months US N = 338 dentate adult aged 65+ blacks (169) and whites (169)	Incidence rates defined as CAL ≥ 3 mm. Extent and severity	Initial periodontal status: no new CAL; CAL in ≥ 1 site previously diseased (baseline CAL < 3 mm); CAL ≥ 1 site previously diseased (baseline CAL ≥ 3 mm); and those who experienced both new disease and progression of existing disease. Sociodemographic, psychosocial, medical, environmental, behavioural and oral information	The 3-year incidence was 41.3% for those with no progression and no new lesions; 27.5% for those with new lesions only; 11.1% for those with progressing lesions only; and 20.1% for those with both. Income < \$15,000, soft-tissue reaction caused by medication, use of smokeless tobacco, and history of pain were significantly associated with an increased incidence of new lesions. Income < \$15,000, soft-tissue reaction caused by medication, smoking cigarettes, BANA positive test, <i>Porphyromonas gingivalis</i> , and financial problems were associated with disease progression
Kaldahl et al. (1996)	To evaluate the effects of the level of cigarette consumption and smoking history on the response to active periodontal treatment and up to 7 years of supportive periodontal treatment. Exams at baseline (Exam 1), 4 weeks after mechanical plaque control instructions (Exam 2), 10 weeks following periodontal surgery (Exam 3), and yearly during 7 years of supportive periodontal treatment. US N = 74, with moderate-to-advanced periodontitis	PD, gingival recession, CAL and horizontal probing attachment level; percentage of sites with plaque and BOP	Smoking status: heavy smokers (≥ 20 cigarettes/day, n = 31); light smokers (≤ 19 cigarettes/day, n = 15); past smokers (had a history of smoking but quit before Exam 1, n = 10), and never smokers (n = 18)	After 7 years of follow-up, past and never smokers consistently exhibited a significantly greater reduction in PD than heavy and light smokers. Past and never smokers exhibited gains in mean CAL compared with light and heavy smokers. There were no differences in average gingival recession of the mean change in horizontal attachment at furcation sites following active therapy among the four groups. All groups experienced a similar decrease in the prevalence of bleeding sites following active therapy
Beck et al. (1997a)	To examine time-to-event analysis of attachment loss over a 5-year period by a variety of characteristics, and to describe a multivariate logistic regression model with potential risk factors. Exams at 18, 36, and 60 months US N = 540 dentate adult aged 65+ (18,947 sites)	Incidence rates defined as CAL ≥ 3 mm	<i>Actinobacillus actinomycetemcomitans</i> , <i>Prevotella intermedia</i> and <i>P. gingivalis</i> . Age, gender, missing teeth, education, smoking, dental visit	In a multivariate model adjusting for the time intervals, smoking (OR = 1.6; 95% CI: 1.2–2.0), being positive for <i>P. gingivalis</i> at baseline (OR = 1.7; 95% CI: 1.3–2.2), having 5 or more missing teeth (OR = 1.9; 95% CI: 1.4–2.7), and not being a high-school graduate (OR = 1.8; 95% CI: 1.4–2.4) were significantly associated with an increased risk for attachment loss. In addition, having a dental visit more than 5 years ago (OR = 1.2; 95% CI: 1.0–1.3) significantly increased the risk

Beck et al. (1997b)	To examine (1) whether attachment loss during one time period is associated with a higher risk for attachment loss at a subsequent period in the same subject; (2) whether sites in survivor teeth with deeper periodontal pockets at baseline are more likely to experience future attachment loss; and (3) whether an effect of regular use of dentists' services on attachment loss are demonstrable in a community-dwelling population. Exams at baseline, 18 months, 36 months, and 5 years US $N=220$ dentate adult aged 65+ blacks (106) and whites (114)	CAL ≥ 3 mm over each 18-month period	PD and dental care	Whites tended to be less likely to experience CAL, with 63.5% experiencing no loss over the 5 years compared with 44.1% of blacks experiencing no loss. Sites with CAL in whites had lower RR of CAL than sites for blacks during all three periods. Baseline pocket depth and irregular dental visits were positively associated with the proportion of sites that demonstrated breakdown over the next 5 years. Sites with PD > 3 mm at baseline consistently associated with higher percentage of sites with CAL, regardless of race
Firathi (1997)	To examine the clinical status of the periodontal tissues in a group of diabetic subjects over 5 years; and to demonstrate the association between DM and periodontal disease. US $N=64$ adolescents (44 with type 1 diabetes and 20 without); whites	Plaque and gingival indices, BOP, PD and CAL	Serum fructosamine, glycosylated haemoglobin and glucose levels	Mean CAL significantly increased between baseline (2.39) and the 5-year follow-up (3.51) in the diabetic group. A significant positive correlation was observed between the duration of diabetes and CAL at both baseline and 5-year examination.
Tervonen and Karjalainen (1997)	To assess variations in periodontal status of type 1 diabetic adolescents; and to study the healing and recurrence of periodontal disease after the hygienic phase of periodontal therapy. Exams at 4 weeks, 6 and 12 months. Finland $N=46$ (36 diabetics and 10 controls) aged 24–36 years	Periodontal status: dental plaque, calculus, PD, BOP and CAL	Diabetic status: D1 ($n=13$) no diabetic complications and good long-term metabolic control; D2 ($n=15$) moderate metabolic control with/without retinopathy; D3 ($n=8$) severe diabetes with poor metabolic control and/or multiple complications	No statistically significant differences in the periodontal health status were observed between the diabetic group as a whole and the non-diabetic controls at any examination. The level of periodontal health of the diabetic patients with good control and no complications (D1) and those with moderate control with/without retinopathy (D2) was similar to non-diabetic controls. Diabetic subjects with poor metabolic control and/or multiple complications (D3) exhibited higher extent of CAL ≥ 2 mm at baseline and higher recurrence of PD ≥ 4 mm during follow-up
Baelum et al. (1997)	To describe predictors of new and progressing destructive periodontal disease over a 10-year period in a rural subject sample. China $N=398$ Aged 20–60+ years	'New disease' was defined as loss of attachment of 2+ mm, in sites with no CAL at baseline; 'progressing disease' was defined as additional attachment loss of 2+ mm at sites with pre-existing attachment loss	Other covariates: age; sex; number of mobile teeth; percentage of sites with plaque; percentage of sites with calculus; percentage of sites with BOP; percentage of sites with baseline of CAL 1+, 4+, and 7+ mm; and percentage of sites with PD of 4+ and 7+ mm	Adjusted analyses showed that male gender, lower # of sites present, increasing percentage of existing sites with CAL 4+ mm and increasing number of mobile teeth were predictors of new disease, while younger age, increasing percentage of existing sites with CAL 4+ mm, and PD with 4+ mm were predictors of PDS

(continued)

Table 19.7 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Papapanou et al. (1997)	In a sample of rural subjects with limited access to dental care, to describe the prevalence of selected bacterial species; correlate the microbiological and clinical profiles of the subjects; and assess the association between the microbiological variables and the retrospective longitudinal changes of periodontal status occurring over a 10-year period. $N=148$, 30–39 and 50–59 years old. China	Tooth site-based analysis: 'deep' sites, if $PD \geq 5$ mm, otherwise 'shallow'; 'progressing' if 10-year longitudinal CAL loss ≥ 3 mm, otherwise 'stable'. Subject-based analysis: 'positive' subject if having ≥ 3 sites with $PD \geq 5$ mm, otherwise 'negative'; 'downhill' if having ≥ 10 sites with 10-year longitudinal CAL loss of ≥ 3 mm, otherwise 'stable'	A maximum of 14 subgingival plaque samples obtained from each subject (1,864 in total), analysed with respect to 18 bacterial species	Ubiquitous prevalence for the majority of the investigated species on the subject level. Bacterial profiles identified 'deep' sites with 74% sensitivity and 64% specificity. Corresponding figures were 66% and 58% for 'progressing' sites; 79% and 66% for 'positive' subjects; and 75% and 85% for 'downhill' subjects. Bacterial colonization at high levels by <i>P. gingivalis</i> , <i>P. intermedia</i> , <i>P. nigrescens</i> , <i>T. forsythensis</i> , <i>F. nucleatum</i> , <i>T. denticola</i> , <i>M. micros</i> , and <i>C. rectus</i> conferred statistically significant ORs for both 'positive' and 'downhill' subjects
Grossi et al. (1997b)	To determine the effect of smoking on the clinical and microbiological response to mechanical periodontal therapy; and to determine the effect of smoking cessation on periodontal therapy. Baseline and 3-month exams US $N=143$ aged 35–65 years with established periodontitis (≥ 2 inter-proximal sites with $CAL \geq 6$ mm and one inter-proximal site with $PD \geq 5$ mm)	PD, CAL, PI, bleeding index	Supragingival scaling and oral hygiene instruction, followed by four to six sessions of subgingival instrumentation. After the first four sessions of subgingival scaling, subjects received one or two additional sessions as needed. Smoking status: current ($n=60$), former ($n=55$) and non-smokers ($n=28$). Overall lifetime exposure to tobacco was quantified in pack-years (# of cigarettes/day $\times$ # of years smoking). Period free or quit years (# of years since smoking was stopped) was also calculated. <i>T. forsythensis</i> and <i>P. gingivalis</i>	There were no differences in mean CAL and PD at baseline by smoking status. After 3 months, there was a significant reduction in whole-mouth mean PD, with current smokers showing less reduction ($p<0.04$) than former- and non-smokers (0.33 ± 0.04 , 0.49 ± 0.06 and 0.49 ± 0.08 , respectively). Current smokers exhibited less CAL gain than former and non-smokers (0.32 ± 0.04 , 0.43 ± 0.06 and 0.39 ± 0.08 , respectively). At baseline, only current smokers showed higher proportion of <i>T. forsythensis</i> ($P<0.04$) and <i>P. gingivalis</i> (NS) compared with former and non-smokers. After treatment, fewer smokers harboured no <i>P. gingivalis</i> ($P<0.008$) or <i>T. forsythensis</i> (NS) compared with former and non-smokers. Among subjects positive for these pathogens, current smokers exhibited the highest proportions.

Taylor et al. (1998a)	To test the hypothesis that the risk for progressive severe alveolar bone loss is greater in subjects with PC type 2 DM, compared with those without type 2 DM or with BC type 2 DM. Two-year study US $N=359$ adults aged 15–57 years (338 diabetes free, 14 with PC and 7 with BC type 2 DM) Native Americans	Alveolar bone loss: the worst bone score in the dentition was categorized as 0%; 1–24%; 25–49%; 50–74%; and $\geq 75\%$ of the root length. Progression of bone loss: difference between baseline worst score and worst score at follow-up	Glycaemic control: BC ($HbA_1c < 9\%$) and PC ($HbA_1c \geq 9\%$). Age, calculus, gingival and plaque indices, time to follow-up, alcohol consumption, smoking, obesity ($BMI > 27$), coronary heart disease, and gender	In multiple logistic regression, PC diabetic subjects were 11 times more likely (95% CI: 2.5–53.3) to have more bone loss and more pronounced bone loss progression than non-diabetic subjects, but there were no such differences between BC and non-diabetic controls. When compared with BC, PC subjects were five times more likely to have more severe alveolar bone loss and more pronounced bone loss progression but this association did not reach statistical significance. Age, time to follow-up, worst bone loss at baseline and calculus index were significant predictors of bone loss progression
Taylor et al. (1998b)	To test the hypothesis that persons with NIDDM have greater risk of more severe alveolar bone loss progression over a 2-year period than those without NIDDM. 2-year study US $N=359$ adults aged 15–57 years (24 subjects with NIDDM) Native Americans	Same as Taylor et al. (1998a)	Diabetes defined by a plasma glucose concentration 200 mg/dl 2 h after a 75-g oral glucose load. Covariates: age, calculus, gingival and plaque indices, time to follow-up, number of teeth, alcohol consumption, smoking, obesity ($BMI > 27$), systolic blood pressure, coronary heart disease, and gender	Compared with non-diabetic subjects, NIDDM patients showed an increasing trend for more bone loss at follow-up. When stratifying by baseline worst bone score, NIDDM patients with 0% score had a greater bone loss progression than their counterparts with a score of 1–24%. NIDDM was associated with a 4.23 times increased risk (95% CI: 1.8–9.9) of change to a worse bone score when compared with non-diabetics. This association was modified by age, with younger adults exhibiting higher risk for alveolar bone loss progression. Age, time to follow-up, worst bone score at baseline and calculus were independent predictors of bone loss progression

(continued)

Table 19.7 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Machtei et al. (1999)	To explore longitudinally a variety of clinical, systemic, and microbiological markers as possible risk factors in subjects with mild periodontal disease. 2–4-year follow-up. US $N=415$ aged 25–75 years whites (95.6%)	PI, gingival index, PD, RAL, CAL, ACH. Progression of disease was measured as annual change in PD, CAL, percentage of losing and gaining sites	<i>A. actinomycetemcomitans</i> , <i>T. forsythensis</i> , <i>C. rectus</i> , <i>P. intermedia</i> , <i>Capnocytophaga</i> species, <i>P. gingivalis</i> , <i>E. saburreum</i> , <i>F. nucleatum</i> . Covariates: age, gender, smoking (current smokers 15.4%), education, income	Approximately 10% of all sites presented for the second visit with attachment loss exceeding 2 mm (4.4% annually), while only 2.2% of all sites exhibited attachment gain (0.9% annually). The percentage of sites losing ACH was 9.4%. Stepwise regression showed that baseline PD, income, presence of <i>capnocytophaga</i> species, smoking, and thyroid disorder were significant predictors of annual change in PD. Past disease history, SES, smoking, presence of <i>T. forsythensis</i> and <i>Capnocytophaga</i> species, dental visits, and several systemic conditions were significant predictors of annual change on CAL and ACH and tooth loss
Norderyd et al. (1999)	To identify risk factors for severe periodontal disease progression in adults. Approximately 16 years of follow-up Sweden $N=361$ aged 20–60 years	Proximal bone level was measured as percentage of the tooth length. The annual periodontal bone loss was calculated based on a mean tooth length of 22 mm for the entire dentition (third molars were excluded). Periodontal disease progression was defined as >20% bone loss at a proximal site	Covariates: age, supragingival plaque, gingival inflammation, and probing pockets 4+ mm	Age, smoking, supragingival plaque, gingival inflammation and PD 4+ mm were associated with an increase of proximal bone loss of >20% in 6+ sites. After adjustments, age, female gender, high income, smoking, and PD 4+ mm were associated with increased proximal bone loss

Bergström et al. (2000a)	To prospectively investigate the influence of smoking over 10 years on the periodontal conditions of a dentally aware population of musicians. Sweden $N=84$; 16 current smokers, 28 former- and 40 non-smokers	Periodontal disease was defined by the percentage of sites with ≥ 4 mm of PD on each occasion ('diseased sites'). Proportion of sites with GB. Bone height: mean percentage of the root length mesially and distally to all teeth excluding third molars	Smoking: number of cigarettes/day, numbers of years, lifetime exposure (cigarette/day $\times$ years). Other covariates: age, PI	The prevalence of diseased sites was 18.7% for current, 11.1% for former-, and 8.7% for non-smokers at baseline. At 10 years, these figures were 41.6%, 7.8%, and 6.6%. For current smokers, the deterioration increased significantly with increasing cigarette consumption, smoking duration, and lifetime exposure. The mean level of bone height at baseline was 80.3% for current-, 80.7% for former- and 85.1% for non-smokers. At 10 years, there was a significant decrease in mean values for current (76.5%) and former smokers (79.6%) when compared with non-smokers (84.1%). There was no significant difference in GB between baseline and follow-up between smoking groups. There was no significant difference between number of teeth present at baseline and follow-up across smoking groups. After adjusting for age, GB, PI, and frequency of diseased sites at baseline, the 10-year change in the frequency of diseased sites was predicted by current smoking ($P=0.039$)
Timmerman et al. (2000)	To assess the impact of clinical and microbiological baseline characteristics on periodontal disease progression over 7 years. Indonesia $N=255$ aged 15–25 years tea labourers with low education	$PDS \geq 1$ site with $CAL \geq 2$ mm	PI, BOP, PD, and CAL score on the buccal aspect of all teeth. <i>A. actinomycetemcomitans</i> , <i>P. gingivalis</i> , <i>P. intermedia</i> , spirochetes, and motile microorganisms. Age and gender	Out of 255 subjects available at baseline, 160 were re-examined at 7 years. At follow-up, all clinical measures with the exception of PD showed a statistical increase that was greater in subjects with PDS at baseline. Presence of <i>A. actinomycetemcomitans</i> (OR = 4.2; 95% CI: 1.4–12.7), <i>P. gingivalis</i> (OR = 2.3; 95% CI: 1.0–5.2) and motile microorganisms (OR = 2.2; 95% CI: 1.0–5.0) were associated with PDS at follow-up. In a multivariable logistic model, age (OR = 1.15, $P=0.04$), subgingival calculus (OR = 1.20, $P=0.02$), and subgingival presence of <i>A. actinomycetemcomitans</i> (OR = 4.61, $P=0.01$) were associated with PDS at follow-up

(continued)

Table 19.7 (continued)

Study	Aim/methods	Outcome(s)	Exposure(s)	Results
Chen et al. (2001)	To examine whether salivary and/or GCF cotinine levels were associated with the severity of periodontal disease in smokers and non-smokers. 10-year study China $N=177$ male farmers, aged 30–69 years	Cotinine in saliva and in GCF. Smoking status: smokers (≥ 2 cigarettes/day, 20 days/month) and non-smokers	PD, CAL, plaque, calculus, and GB. Other covariates: age, packs smoked per year, years smoking	Cigarette smoking was associated with a greater increase in PD and attachment loss, as well as greater tooth loss at an earlier age. All smokers had detectable salivary and GCF cotinine. However, neither salivary cotinine nor GCF cotinine was significantly correlated with PD, attachment loss, and tooth loss ($P>0.05$)
Stewart et al. (2001)	To compare changes in glycaemic control in a group of patients with type 2 DM following periodontal treatment, to a control group with type 2 DM who did not receive periodontal treatment. 18 months US $N=72$, equally divided between treatment and control groups	Periodontal treatment group: root planing, subgingival curettage and extractions. Control group: no dental treatment	Glucose control: % HbA1c. Covariates: age, race	Patients receiving periodontal treatment showed a decrease of 17.1% ($P=0.0001$) in HbA1c levels, while the control group showed a decrease of 6.7% ($P=0.02$). Changes in HbA1c were statistically significant for those receiving oral hyperglycaemic agents (18%, $P=0.0005$) and insulin (18%, $P=0.003$) in the treatment group only
Cullinan et al. (2001)	To investigate the relationship between IL-1 genotype and periodontitis. 5-year study Australia $N=295$	PD and CAL at 6, 12, 24, 36, 48, and 60 months. Periodontitis progression: CAL ≥ 2 mm at any period of exam during the 5 years	SNPs at IL-1A (+4845 T) and IL-1B (+3954 T). <i>A. actinomycetemcomitans</i> , <i>P. intermedia</i> , and <i>P. gingivalis</i> . Age, gender, smoking	The prevalence of the composite genotype consisting of the less common allelic variants was 39%. A relationship was found between the IL-1 positive genotype and increased mean probing pocket depth in non-smokers >50 years old. IL-1 genotype-positive smokers and genotype-positive subjects with <i>P. gingivalis</i> had a higher number of sites with PD ≥ 3.5 mm. There was a trend for IL-1 genotype-positive subjects to experience higher periodontitis progression than IL-1 genotype-negative subjects

Ogawa et al. (2002)	To identify risk factors for periodontal disease progression among elderly people over a 2-year period. Japan <i>N</i> = 554 (281 males and 273 females) at baseline, <i>N</i> = 394 (208 males and 186 females) at follow-up	Periodontal disease progression defined as ≥1 site with ≥3 mm longitudinal CAL between the two examinations	Covariates: gender, smoking, alcohol, use of dental care, dental visits per year, dental self-care behaviours, CAL ≥ 6 mm, and number of teeth present. Blood pressure, serum markers of liver disease (GOT, GPT and γ-GTP), kidney disease (creatinine), immunoglobulins (IgG, IgA and IgM), lipid profiles (total cholesterol and triglycerides), and nutritional factors (total-protein, calcium, blood-sugar, and albumin)	Seventy-one percent of the study population exhibited an additional CAL 3+ mm at one or more sites after 2 years. Males, smokers, and subjects with CAL 6+ mm at baseline were more likely to lose an additional 3+ mm of CAL. In the multivariable analyses, smoking, gender, CAL ≥ 6 mm, and ≤20 remaining teeth were associated with disease progression
Rodrigues et al. (2003)	To compare changes in glycaemic control in a group of patients with type 2 DM following full-mouth scaling and root planing alone or in combination with amoxicillin/clavulanic acid. 3-month follow-up. Brazil <i>N</i> = 30 with chronic periodontitis	Chronic periodontitis: at least 1 site with PD ≥ 5 mm and 2 teeth with CAL ≥ 6 mm	Subjects randomly assigned to full-mouth scaling and root planing alone (<i>N</i> = 15; G2) and in combination with amoxicillin/clavulanic acid (<i>N</i> = 15; G1). HbA1c and fasting glucose levels assessed at baseline and 3 months	HbA1c values were reduced for both groups between baseline and 3 months. However, only the changes in G2 were statistically significant ($8.8 \pm 1.8\%$ versus $7.6 \pm 1.4\%$, $P = <0.05$). Baseline mean fasting glucose levels for both groups were statistically different ($P < 0.05$); G1: 221 ± 60 mg/dl and G2: 175 ± 68 mg/dl. There was no difference between groups in the mean fasting glucose levels at 3 months

CAL clinical attachment level, *RR* relative risk, *OR* odds ratio, *CI* confidence interval, *PD* probing depth, *PC* poorly controlled, *DM* diabetes mellitus, *BC* better controlled, *NIDDM* non-insulin-dependent diabetes mellitus, *BMI* body mass index, *RAL* relative attachment level, *ACH* alveolar crestal height, *SES* socioeconomic status, *PDS* progressive disease, *PI* plaque index, *BOP* bleeding on probing, *GCF* gingival crevicular fluid, *GB* gingival bleeding, *IL* interleukin, *SNP* single-nucleotide polymorphism, *HbA1c* glycated haemoglobin

Table 19.8 Adjusted odds ratios (ORs) of the association between different periodontal measurements (most of them continuous) and adverse pregnancy outcomes (only statistically significant results shown) (Manau et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Clinical examination	Clinical measurement	PTB ^a OR (95% CI)	LBW ^b OR (95% CI)	PLBW ^c OR (95% CI)	PTB and/or LBW ^d OR (95% CI)
Full mouth six sites per tooth	% sites CAL ≥ 3 mm	1.00 (1.00–1.02)			
	% sites CAL ≥ 4 mm	1.02 (1.00–1.03)		1.51 (1.11–2.10)	1.51 (1.11–2.10)
	% teeth CAL ≥ 4 mm	1.01 (1.00–1.02)		1.01 (1.00–1.02)	1.01 (1.00–1.02)
	Severity CAL ≥ 2 mm	1.51 (1.11–2.10)		1.84 (1.03–3.28)	
	Severity CAL ≥ 3 mm			2.21 (1.02–4.81)	
	Severity CAL ≥ 4 mm				
Full mouth four sites per tooth	% sites CAL ≥ 3 mm	1.01 (1.00–1.02)			1.02 (1.00–1.03)
	% sites CAL ≥ 4 mm	1.02 (1.00–1.03)			
	Severity CAL 1 mm ^e	1.54 (1.12–2.14)			
	Severity CAL ≥ 2 mm	1.51 (1.11–2.10)		1.91 (1.03–3.55)	
	Severity CAL ≥ 3 mm			2.34 (1.03–5.28)	
	Severity CAL ≥ 4 mm			1.01 (1.00–1.02)	1.01 (1.00–1.02)
Full mouth sextants	% sextants CAL > 3 mm	1.01 (1.00–1.02)			
	% sextants CAL > 5 mm	1.02 (1.00–1.03)		1.02 (1.00–1.03)	
	% sites CAL ≥ 3 mm	1.00 (1.00–1.02)			
Partial recording three buccal sites/tooth two diametrically opposed quadrant	% sites CAL ≥ 3 mm	1.01 (1.00–1.02)			
	% sites CAL ≥ 3 mm	1.01 (1.00–1.02)			
	% sites CAL ≥ 3 mm	1.00 (1.00–1.01)			

CAL clinical attachment loss, PTB pre-term birth, LBW low birth weight, PLBW premature low birth weight, CI confidence interval

^aOR adjusted by: Age, place of residence, race, smoking habit, systemic disease, previous prematurity, spontaneous miscarriage, number of pregnancies, onset of prenatal care, complications of pregnancy type of delivery, untreated dental decay and plaque index

^bOR adjusted by: Age, race, smoking habit, systemic disease, BMI, previous low birth weight, spontaneous miscarriage, number of pregnancies, onset of prenatal care, complications of pregnancy, type of delivery, untreated dental decay and time since last oral prophylaxis

^cOR adjusted by: Age, race, smoking habit, systemic disease, X-rays during pregnancy, previous low birth weight, previous prematurity, spontaneous miscarriage, number of pregnancies, onset of prenatal care, complications of pregnancy, gestational diabetes and type of delivery

^dOR adjusted by: Age, place of residence, race, smoking habit, systemic disease, BMI, previous low birth weight, previous prematurity, spontaneous miscarriage, number of pregnancies, onset of prenatal care, complications of pregnancy and type of delivery

^eSeverity index: mean CAL in excess of 1 mm for sites with CAL > 1 mm

least four teeth. ‘PD4–2T periodontitis’ was defined as the presence of one or more sites with a PD of ≥ 4 mm in at least two teeth. A PD of ≥ 4 mm has been suggested as a threshold for evaluating treatment needs. ‘CAL3–2T periodontitis’ was defined as the presence of one or more proximal sites with a CAL of ≥ 3 mm in at least two teeth that might be adjacent. ‘PD5 or CAL4–2T periodontitis’ was defined as the presence of at least two proximal sites with a PD of ≥ 5 mm (not on the same tooth) or at least two proximal sites with a CAL of ≥ 4 mm (not on the same tooth). ‘PD4–CAL3–2T periodontitis’ was defined as the presence of at least two teeth with sites with both PD ≥ 4 mm and CAL ≥ 3 mm at the same site. To determine the extent of the disease, ‘PD4–CAL3–4T periodontitis’ was defined as the presence of sites with both PD ≥ 4 mm and CAL ≥ 3 mm in at least four teeth, which corresponds to approximately 30% of affected teeth. ‘PD4–CAL3–BoP–2T periodontitis’ was defined as the presence of sites with a PD of ≥ 4 mm, a CAL of ≥ 3 mm, and BoP at the same site in at least two teeth. It was showed that the definition of periodontitis had an impact on the frequency of periodontitis, which ranged from 12.1% to 37.7%, and produced different ORs for the associations with risk factors for periodontitis.

Different case definitions can have also a great impact on the prevalence and extent rates of periodontitis. In this manner, it can influence the results and associations presented in studies as well as over or underestimate the real need for periodontal treatment (Costa et al. 2009).

The accuracy of prevalence estimates for periodontitis by comparing the ‘true’ prevalence computed from the ‘gold standard’ full-mouth periodontal examination data with the prevalence determined from the NHANES III and 2001–2004 partial-mouth periodontal examination (PMPE) protocols, respectively, while using the same collected data was determined (Eke et al. 2010). Both these NHANES protocols specified the examination of each tooth (excluding third molars) in two randomly selected quadrants (half jaws), namely, one upper (maxillary) and one lower (mandibular). The NHANES III protocol specified 2 sites per tooth (mesio- and mid-buccal)

involving up to 28 sites, whereas 3 fixed sites per tooth (mesio-, mid-, and disto-buccal) encompassed up to 42 sites per individual according to the NHANES 2001–2004 protocol.

For the comparisons, the following seven different periodontitis case definitions were applied, all previously used in official US public health publications on national prevalence of periodontitis (Eke et al. 2010):

1. CDC/AAP (American Academy of Periodontology) (hereafter referred to as ‘CDC’) definition based on measures of CAL and PD at interproximal sites (Page and Eke 2007): *Severe periodontitis* is defined as having two or more interproximal sites with ≥ 6 -mm CAL (not on the same tooth) and one or more interproximal site(s) with ≥ 5 -mm PD. *Moderate periodontitis* is defined as two or more interproximal sites with ≥ 4 -mm clinical CAL (not on the same tooth) or two or more interproximal sites with PD ≥ 5 mm, also not on the same tooth. *Total periodontitis* is the sum of severe and moderate disease.
2. Health and Human Services Vital and Health Statistics Series 11 Report (‘Series 11’) case definition requires at least one site with ≥ 3 -mm CAL and ≥ 4 -mm PD (Arbes et al. 2001).
3. The Healthy People 2010 (‘HP 2010h’) operational definition found periodontitis with CAL ≥ 4 mm at two or more sites (US Department of Health and Human Services 2000). Additionally, four dichotomized case definitions at least one site(s) with (4) ≥ 6 -mm CAL, (5) ≥ 3 -mm CAL, (6) ≥ 7 -mm PD, and (7) ≥ 5 -mm PD.

According to the CDC case definitions, the total true prevalence of periodontitis from the ‘gold standard’ FMPE protocol was 22.4%, namely, 4.8% severe and 17.5% moderate periodontitis. Prevalence of disease according to Series 11 and HP 2010 case definitions was 34.4% and 30.3%, respectively. Prevalence assessed by at least one site with CAL ≥ 6 mm was 13.6% and increased fourfold to 63.2% with the lower threshold of ≥ 3 -mm CAL. Similarly, periodontitis prevalence expressed by site(s) with 7+ mm PD was 3.7% but became four times greater (15.4%) with 5+ mm (Table 19.9). When the CDC definitions were applied to the NHANES

III protocol, a total prevalence of 8.9% was detected versus the true prevalence, determined to be 22.4%, thus underestimating the total true prevalence by 13.4% or about 60% of the true prevalence (relative bias). Prevalence of severe and moderate periodontitis was underestimated by 63% and 59%, respectively (Eke et al. 2010).

Similar findings pertained to use of the NHANES 2001–2004 protocol with the CDC definitions. The highest accuracies were achieved when we used the case definitions of 3- or 6-mm CAL at one or more sites in conjunction with either of the NHANES PMPE protocols. However, the underestimation of true prevalence was still around 40% with the NHANES III and 30% with the NHANES 2001–2004 protocols. The lowest accuracies were attained when case definitions of 5-mm PD were used, in which case 71.4% disease was missed with the NHANES III and 55.7% with the NHANES 2001–2004 protocol. With the latter, the CDC definition identified a total periodontitis prevalence of 9.7%, thus underestimating the true prevalence by about 57% (relative bias). Severe and moderate periodontitis were underestimated by 50% and roughly 60%, respectively. In both NHANES protocols, the accuracies of case definitions requiring combined measures of CAL and PD (e.g., CDC-Severe and Series 11 definitions) were consistently detecting less than 50% of true cases (Table 19.9) (Eke et al. 2010).

The impact of case definition on the prevalence and extent rates of periodontitis was also assessed (Costa et al. 2009). A data set including 340 periodontal records, collected in Belo Horizonte, Brazil, was used. Periodontitis was defined as (1) one site with probing depth (PD) ≥ 4 mm; (2) clinical attachment level (CAL) ≥ 5 mm in ≥ 4 sites + one site with PD ≥ 4 mm; (3) CAL ≥ 6 mm in ≥ 2 teeth + one site with PD ≥ 5 mm; (4) $> \text{or} = 4$ teeth with ≥ 1 sites with PD ≥ 4 mm + CAL ≥ 3 mm; (5a) interproximal CAL or PD ≥ 4 mm at ≥ 2 sites, not on the same tooth; and (5b) interproximal CAL of ≥ 6 mm at ≥ 2 sites, not on the same tooth + PD ≥ 5 mm at ≥ 1 proximal site. Definition 5 was determined to be the gold standard, and the definitions were compared by means of agreement, sensitivity, specificity, and positive and negative predictive values. Prevalence and extent

rates greatly varied, from 13.8% to 65.3% and from 9.7% to 55.6%, respectively. As can be concluded from the results, the use of different case definitions has a great impact on the prevalence and extent rates of periodontitis.

19.4.3 Proposal of Definitions of a Periodontitis Case and Disease Progression

Although the problems regarding the compromised validity, comparability, and transferability of the research results generated using different definitions of periodontitis cases have long been acknowledged (Brown and L  e 1993; Papapanou 1996; Baelum and Lopez 2003, 2012; Lopez and Baelum 2003; Tonetti and Claffey 2005; Page and Eke 2007; Manau et al. 2008; Meisel and Kocher 2009; Preshaw 2009; Savage et al. 2009), the attempts to alleviate them have so far not been successful.

Van der Velden (2000) proposed a classification system for periodontitis based on a combination of parameters used in the following order: (1) extent category, (2) severity category, (3) diagnosis on the basis of clinical characteristics if applicable, and (4) diagnosis on the basis of age. The amended version of this classification has a few additions (Van der Velden 2005):

- Defining when periodontitis is considered to be present. It is suggested to define periodontitis as the presence of inflamed pathological pockets ≥ 4 mm deep in conjunction with attachment loss. If present, then the next steps can be taken.
- Classification based on extent of the disease, that is, number of affected teeth.
- Classification based on severity of disease per tooth. The fact that either attachment loss or bone loss can be used for the classification of severity implies that although it may be important to know the actual root length in a given patient, radiographs are not a prerequisite for the classification of severity.
- Classification based on age.
- Classification based on clinical characteristics.

Table 19.9 Accuracy of NHANES periodontal examination protocols in detecting the true prevalence of periodontitis by case definitions (Eke et al. 2010) (Reprinted with permission from Sage Publications)

Examination protocol	Case definition	Periodontitis prevalence (%)	Absolute bias (%) ^a	Relative bias (%) ^b	Sensitivity ^c	Inflation factor
Full-mouth	(1) CDC-severe	4.8	Ref	Ref	Ref	Ref
Full-mouth	(1) CDC-moderate	17.5	Ref	Ref	Ref	Ref
Full-mouth	(1) CDC-total	22.4	Ref	Ref	Ref	Ref
Full-mouth	(2) Series 11	34.4	Ref	Ref	Ref	Ref
Full-mouth	(3) HP2010	30.3	Ref	Ref	Ref	Ref
Full-mouth	(-) ≥1 site with ≥7 mm CAL	8.9	N/A	N/A	N/A	N/A
Full-mouth	(4) ≥1 site with ≥6 mm CAL	13.6	Ref	Ref	Ref	Ref
Full-mouth	(-) ≥1 site with ≥5 mm CAL	22.4	N/A	N/A	N/A	N/A
Full-mouth	(-) ≥1 site with ≥4 mm CAL	41.5	N/A	N/A	N/A	N/A
Full-mouth	(5) ≥1 site with ≥3 mm CAL	63.2	Ref	Ref	Ref	Ref
Full-mouth	(6) ≥1 site with ≥7 mm PD	3.7	Ref	Ref	Ref	Ref
Full-mouth	(-) ≥1 site with ≥6 mm PD	8.1	N/A	N/A	N/A	N/A
Full-mouth	(7) ≥1 site with ≥5 mm PD	15.4	Ref	Ref	Ref	Ref
Full-mouth	(-) ≥1 site with ≥4 mm PD	35.9	N/A	N/A	N/A	N/A
Full-mouth	(-) ≥1 site with ≥3 mm PD	67.3	N/A	N/A	N/A	N/A
NHANES 2001	(1) CDC-severe	2.4	-2.4	-50.2	0.5	2.0
NHANES 2001	(1) CDC-moderate	7.2	-10.3	-58.7	0.4	2.5
NHANES 2001	(1) CDC-total	9.7	-12.7	-56.9	0.4	2.5
NHANES 2001	(2) Series 11	16.9	-17.5	-50.1	0.5	2.0
NHANES 2001	(3) HP2010	16.9	-13.4	-44.2	0.6	1.7
NHANES 2001	(4) ≥1 site with ≥6 mm CAL	9.2	-4.4	-32.3	0.7	1.4
NHANES 2001	(5) ≥1 site with ≥3 mm CAL	44.5	-18.6	-29.5	0.7	1.4
NHANES 2001	(6) ≥1 site with ≥7 mm PD	1.9	-1.8	-47.2	0.5	2.0
NHANES 2001	(7) ≥1 site with ≥5 mm PD	6.8	-8.6	-55.7	0.4	2.5
NHANES III	(1) CDC-Severe	1.8	-3.1	-63.4	0.4	2.7
NHANES III	(1) CDC-Moderate	7.2	-10.3	-58.7	0.4	2.4
NHANES III	(1) CDC-Total	8.9	-13.4	-59.8	0.4	2.4
NHANES III	(2) Series 11	12.8	-21.7	-62.9	0.4	2.4
NHANES III	(3) HP2010	14.9	-15.4	-50.7	0.5	2.0
NHANES III	(4) ≥1 site with ≥6 mm CAL	8.1	-5.5	-40.4	0.6	1.7
NHANES III	(5) ≥1 site with ≥3 mm CAL	39.5	-23.7	-37.6	0.6	1.7
NHANES III	(6) ≥1 site with ≥7 mm PD	1.3	-2.4	-64.6	0.4	2.5
NHANES III	(7) ≥1 site with ≥5 mm PD	4.4	-10.9	-71.4	0.3	3.3

Case definitions for periodontitis: (1) CDC-severe: ≥2 interproximal sites with ≥6 mm CAL (not on the same tooth) and ≥1 interproximal site(s) with ≥5 mm PD. (1) CDC-moderate: ≥2 interproximal sites with ≥4 mm clinical CAL (not on the same tooth) or ≥2 interproximal sites with PD ≥5 mm (not on the same tooth). (1) CDC-total = 'CDC-severe' plus 'CDC-moderate'; (2) Series 11: ≥1 site with ≥3 mm CAL and ≥4 mm PD; (3) HP 2010: ≥2 sites with CAL ≥4 mm. Additionally: ≥1 site with: (4) ≥6 mm CAL; (5) ≥3 mm CAL; (6) ≥7 mm PD; and (7) ≥5 mm PD. For FMPE: (-) Additional, un-numbered case definitions added for comparison: ≥1 site with: ≥7 mm CAL; ≥5 mm CAL; ≥4 mm CAL; ≥6 mm PD; ≥4 mm PD; and ≥3 mm PD

Full-mouth FMPE, NHANES 2001 NHANES 2001–2004, PD pocket depth, CAL clinical attachment loss, CDC CDC/AAP

^aAbsolute bias = Absolute difference = (PMPE-protocol prevalence) – (true prevalence determined by FMPE)

^bRelative bias = Percent underestimation of prevalence = percent of absolute difference relative to true prevalence: [(absolute difference)/(true prevalence)] × 100

^cSensitivity = [(PMPE-protocol prevalence)/(true prevalence)] × 100

The classification is ascertained in the following way: (a) first, the severity category is determined for each tooth; (b) next, the extent category is determined by counting the number of teeth with the most severe condition; (c) diagnosis on the basis of clinical characteristics is added if applicable; and (d) diagnosis on the basis of age (Van der Velden 2005).

In Europe, at the 2005 *European Workshop on Periodontology*, it was stated that: based on the existing literature on risk factors for periodontitis, it is not possible to ascertain whether the reported variance in odds ratios or relative risk is because of a varying biological impact of the factor under investigation in different populations or merely reflects the inconsistency in case definitions or thresholds for progression used across studies. In order to establish a framework that allows some consistency of data interpretation across global epidemiological studies, it is necessary for all studies to use one consistent definition for “a periodontitis case” and “periodontitis progression”. Such an approach allows odds ratios and estimates of relative risk, both of which are sensitive to threshold definition, to be derived that are directly comparable between different studies. (Tonetti and Claffey 2005).

In this context, it is important to recognize that:

1. Periodontitis cannot be reflected by measurements of a single variable. While past experience of periodontitis is reflected by attachment loss/bone loss measurements, assessment of disease presence requires additional measurement of bleeding on probing and/or probing pocket depth.
2. The accepted measure of cumulative lifetime experience of periodontitis is attachment loss; therefore, this measure should be the primary outcome variable used in studies of risk factors for periodontitis (Tonetti and Claffey 2005).

Proposed criteria for a two-level periodontitis case definition by Tonetti and Claffey (2005):

1. Presence of proximal attachment loss of ≥ 3 mm in ≥ 2 non-adjacent teeth
2. Presence of proximal attachment loss of ≥ 5 mm in $\geq 30\%$ of teeth present

Rationale for definition of periodontitis case:

- Two threshold levels are proposed: the first to enable the utilization of a sensitive case definition

(inclusive of incipient cases) and the second to allow a more specific case definition (to identify only cases with substantial extent and severity).

- Proximal sites and non-adjacent teeth are specified, in order to minimize the likelihood of including attachment loss affecting buccal/lingual sites or adjacent interdental sites for reasons other than periodontitis.
- The 3-mm threshold is based upon studies of incremental attachment loss measurement, where the error of the recording method was calculated at 2.5 mm.
- The proposed criteria are not designed for the assessment of prevalence of periodontitis across populations and/or age groups; the focus is to identify risk factors.

Proposed criteria for periodontitis progression case definition by Tonetti and Claffey (2005):

Presence of ≥ 2 teeth demonstrating a longitudinal loss of proximal attachment of ≥ 3 mm. In situations where serial proximal attachment level measurements are not available, longitudinal radiographic bone loss of ≥ 2 mm at ≥ 2 teeth may be used as a substitute.

Rationale for definition of a case of periodontitis progression:

- The threshold for periodontitis progression is based upon extensively documented evidence within the periodontal literature.
- The threshold is set at the level of two teeth to minimize the risk of including cases of progression arising because of reasons other than periodontitis (Tonetti and Claffey 2005).

The American Academy of Periodontology report stated that determining the prevalence of periodontitis in the US population, seemingly a straightforward issue, in fact is complicated by the various case definitions used. If periodontitis is defined as the identification of at least one site with CAL of ≥ 2 mm, around 80% of all adults are affected and around 90% of those aged 55–64. When the case definition is at least one site with CAL of ≥ 4 mm, the prevalence in those aged 55–64 drops to around 50%. When it is CAL of ≥ 6 mm, prevalence is less than 20%. Using pockets of ≥ 4 mm as a case definition, 30% of adults had met that criterion on at least three to four teeth (Burt 2005).

Finally, in February 2003, the Division of Oral Health at the Centers for Disease Control and Prevention (CDC), in collaboration with the American Academy of Periodontology (AAP), appointed a working group to examine the feasibility of, and to identify valid nonclinical measures for, population-based surveillance of periodontitis. The case definitions for periodontitis developed by the CDC Periodontal Disease Surveillance Workgroup are as follows. Two definitions are provided for periodontitis, one for severe periodontitis and another for moderate periodontitis (Page and Eke 2007):

- The case definition for **severe periodontitis** requires two or more interproximal sites with $CAL \geq 6$ mm, not on the same tooth, and one or more interproximal sites with $PD \geq 5$ mm. Interproximal sites, in contrast to buccal or lingual sites, are required because the disease usually begins and is most severe at interproximal sites and because this minimizes the effects of gingival recession on the accuracy of the PD measurements. At least two sites with $CAL \geq 6$ mm, not on the same tooth, are required because it is possible to have abnormal CAL and not have periodontitis. Such conditions include a subgingival restoration with an overhanging margin and the distal aspect of some mandibular second molars where a third molar has been extracted. In addition, the requirement takes into account evaluator variation and the underestimation of disease known to result from partial-mouth examinations (Page and Eke 2007).
- **Moderate periodontitis** was defined as two or more interproximal sites with $CAL \geq 4$ mm, not on the same tooth, or two or more interproximal sites with $PD \geq 5$ mm, not on the same tooth. As with the definition for severe periodontitis, the case definition for moderate periodontitis includes two or more affected sites and sites with CAL, abnormal PD, or both (Page and Eke 2007).

It is hoped that these definitions will serve as standard case definitions for population-based surveillance of moderate and severe periodontal disease for the future, which will bring some uniformity to case definitions of the disease across studies (Page and Eke 2007).

Recently, the extent to which the three periodontitis case definition systems proposed by van der Velden, Tonetti and Claffey (TC), and Page and Eke (PE) identify the same cases in a population of never-treated adults with limited tradition for oral hygiene procedures was assessed. Based on data on clinical attachment level (CAL), probing pocket depth (PD), and bleeding on probing (BOP) of four sites in all teeth present among 1,130 adult Kenyans, the population was classified according to the three case definition systems, and according to the occurrence of the concomitant presence of CAL and BOP at the site level (Baelum and Lopez 2012). When comparing the diagnoses obtained with the TC and PE case definitions, 12% of the subjects were misclassified, that is, identified as a periodontitis case by one but not by the other case definition system (Table 19.10). The authors concluded that the classification system proposed by van der Velden is better suited for providing clinicians with a clear image of the case. As Baelum and Lopez (2012) showed the three classification/case definition systems evaluated herein have been constructed for three different purposes. The criteria proposed by Tonetti and Claffey on behalf of the European Workshop in Periodontology (Tonetti and Claffey 2005) were intended for research purposes, while the criteria proposed by Page and Eke (2007) were motivated by population surveillance concerns. Originally, the classification system proposed by van der Velden was constructed to provide a classification system that is ‘not susceptible to multiple interpretations’ and classifies all patients with periodontitis (van der Velden 2000). The later version of the system aims to describe the patients ‘in such a way that all clinicians immediately have a clear image of a case’ (van der Velden 2005). However, Baelum and Lopez (2012) contested the idea that different purposes need different sets of case-defining criteria. At the end of the day, most clinical research is carried out with the view to ultimately translate the research results into clinical practice, and the transferability of results from research to clinical practice would be greatly facilitated if the classifications used in research did actually correspond to those used in clinical

Table 19.10 Diagnostic performance of the periodontitis case definitions proposed by Tonetti and Claffey (2005) and Page and Eke (2007), respectively, for the correct identification of subject with different periodontal conditions (Baelum and López, 2012)(Reprinted with permission from John Wiley & Sons)

Gold Standard definition of periodontitis cases	Population prevalence (%)	Diagnostic test (case definition)													
		Tonetti and Claffey sensitive+specific				Page and Eke moderate+severe				Tonetti and Claffey specific only					
		Sens.	Spec.	Mis-class. (%)		Sens.	Spec.	Mis-class. (%)		Sens.	Spec.	Mis-class. (%)			
Page and Eke – mod- erate + severe cases	63	0.93	0.78	12		–	–	–		–	–	–	–		–
Tonetti and Claffey – sensitive + specific cases	67	–	–	–		0.88	0.87	12		–	–	–	–		–
Page and Eke – severe cases only	29	–	–	–		–	–	–		0.69	0.96	12	–		–
Tonetti and Claffey – specific cases only	22	–	–	–		–	–	–		–	–	–	0.88	0.89	12
≥1 site with CAL ≥1 mm & BOP	89	0.75	0.98	23		0.70	0.92	28		–	–	–	–	–	–
≥10 sites with CAL ≥1 mm & BOP	69	0.91	0.89	9		0.83	0.82	17		0.32	1.00	47	0.41	0.99	42
≥30 sites with CAL ≥1 mm & BOP	49	0.99	0.64	19		0.92	0.65	22		0.42	0.97	30	0.51	0.93	27
≥1 site with CAL ≥4 mm & BOP	70	0.91	0.90	9		0.85	0.97	15		0.32	1.00	47	0.41	1.00	41
≥10 sites with CAL ≥4 mm & BOP	41	1.00	0.56	26		0.99	0.62	23		0.54	1.00	19	0.65	0.97	16
≥30 sites with CAL ≥4 mm & BOP	20	1.00	0.42	47		1.00	0.46	43		0.82	0.93	9	0.83	0.85	15

(–) Comparison not relevant
CAL clinical attachment level, BOP bleeding on probing

settings. Similarly, there would be little point in the use of a special classification system for surveillance purposes if the results of the surveillance could not be translated into needs and demands at the clinical level.

19.5 Periodontal Disease Progression

A common problem in periodontal research is the complexity of clinical data since the assumption underpinning most statistical methods is that observations are independent. However, data are usually collected on several different occasions, at several sites around different teeth, within several subjects (Gilthorpe et al. 2003). Aggregation of information to the subject level has led to the impression that progression of periodontal diseases is a continuous phenomenon, occurring at a relatively constant pace, only varying between population groups (Löe et al. 1978; Gilthorpe et al. 2003).

However, in the mid-1980s, data from longitudinal monitoring of periodontal attachment levels and alveolar bone in humans and in animals suggested that periodontal disease progresses by recurrent acute episodes. In addition, rates of attachment loss have been measured in individual sites which are faster than those consistent with the continuous disease hypothesis or slower than those expected from estimates of prior loss rates. To account for these observations, a model of destructive periodontal disease is described in which bursts of activity occur for short periods of time in individual sites. These bursts appear to occur randomly at periodontal sites throughout the mouth. Some sites demonstrate a brief active burst of destructive periodontal disease (which could take a few days to a few months) before going into a period of remission. Other sites appear to be free of destructive periodontal disease throughout the individual's life. The sites which demonstrate destructive periodontal activity may show no further activity or could be subject to one or more bursts of activity at later time periods. Comparison of monitored loss rates for a year with mean loss rates prior to monitoring

suggested that there may be relatively short periods in an individual's life in which many sites undergo periodontal destruction followed by periods of extended remission. An extension of the random disease model is also suggested in which bursts of destructive periodontal disease activity occur with higher frequency during certain periods of an individual's life (Socransky et al. 1984).

This gave rise to two competing theories: the **'linear' (continuous-rate) model**, where, overall, sites slowly and progressively lose attachment and the **(random) 'burst' model**, where multiple sites show breakdown within a short period, with periods of remission that might last for months, years, or decades (Gilthorpe et al. 2003).

It has been argued that measurement error, and the possibility of different disease patterns at different sites, could produce erroneous evidence of burst progression (Gilthorpe et al. 2003). It was suggested that bursts of 2 mm or smaller cannot be reliably distinguished from linear disease progression using the lack-of-fit test, except under unusual clinical circumstances. Under typical clinical circumstances, burst sizes needed to be 3–5 mm in order to be reliably distinguished from linear disease progression (Hujóel and Leroux 1998).

In a meta-analysis performed in 1999, it was concluded that few papers have addressed the question of site-specific attachment level change in untreated chronic adult periodontitis. Only 8 out of 23 papers reviewed from 1982 to 1997 have adequate data; only one paper (Haffajee et al. 1983) described losing sites as well as gaining sites and sites showing exacerbation/remission patterns of change. There are considerable differences in the probes used, in the thresholds achieved with these probes, in the number of measurements taken, in the number of subjects and sites studied, and in the study length. There are also considerable differences in the proportions of sites showing attachment loss, gain, and exacerbation/remission patterns between individual studies and between the probe generations (Breen et al. 1999a).

In this context, it must also be realized that, as long as disease progression is measured by linear

measurements of vertical attachment loss along the root surface, 'bursts' of activity will be the de facto favoured alternative because the magnitude of the detectable progression is directly dependent on the incremental readings of the periodontal probe (Borrell and Papapanou 2005). Jeffcoat and Reddy (1991) monitored 30 patients with adult periodontitis using an automated periodontal probe for 6 months in order to determine the prevalence of active sites and the overall pattern of disease progression in active sites. The automated probe is capable of measuring probing attachment levels relative to the cemento-enamel junction with better than 0.2 mm of accuracy. The prevalence of disease activity was dependent on the threshold for probing attachment loss used. When the smallest threshold (0.4 mm) was used, the prevalence of active disease was 29%, whereas a large threshold (2.4 mm) detected only 2% active sites. Regression analysis of the active sites revealed that 76% of the sites lost probing attachment consistent with a continuous model for disease progression. A small subset of sites demonstrated either bursts of activity or exacerbations and remissions of disease activity.

The option-4 algorithm was designed to reduce measurement error and improve accuracy and sensitivity of longitudinal site-specific attachment level change (SSAC) detection. The algorithm allows a maximum of four site-specific measurements per visit with within- and between-visit comparisons; the threshold used for these comparisons is 1.0 mm. A precalibrated clinician recorded full-mouth RAL with a third-generation disc probe on four occasions over 6 months in 16 subjects (mean age 48.1 years) with moderately advanced chronic adult periodontitis (2,312 sites). SSAC was detected in 100% subjects and at 51.0% measured sites. Linear SSAC ($R^2 \geq 0.90$; $P \leq 0.05$) occurred at 105 sites (4.5%): 32 sites (1.4%) deteriorated and 73 sites (3.1%) improved. Between-visit SSAC occurred at 1,074 sites (46.5%): 391 sites (16.9%) deteriorated, 295 sites (12.8%) improved, and 388 sites (16.8%) showed exacerbation/remission patterns (Breen et al. 1999b). The option-4 algorithm produced a reduction of 75.6% in the mean site-specific variance of probeable crevice depth measurements

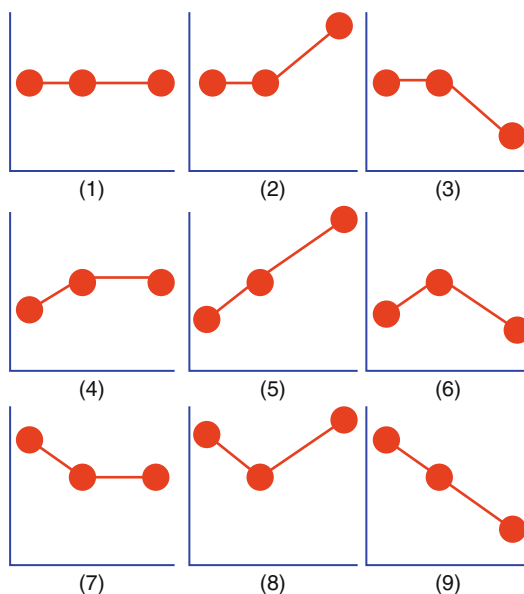


Fig. 19.1 A summary of all the possible changes in either periodontal outcome across the study period. The *nine charts* represent all possible transitions across three time points, where the *red* represent hypothetical outcome values at each measurement occasion. Outcomes may remain constant (chart 1) or after no initial change, may exhibit change from occasions two to three (charts 2, 3) or after initial progression, there may be no change (chart 4), further progression (chart 5), or improvement (chart 6), and similarly, following initial improvement (charts 7–9) (Gilthorpe et al. 2003. Reprinted with permission from Sage Publications)

(PCD) PCD1/PCD2 and a 75% reduction in the median site-specific variance of PCD1/PCD2 (Breen et al. 1999c).

Multilevel modelling provides a new approach, yielding a more comprehensive model. In brief, a multilevel model is a regression model for data that form a hierarchy, with variation at each level determined separately. Coefficients for explanatory variables may exhibit random variation about their mean estimate (at any level), leading to the inclusion of additional variance and covariance terms. Each covariance provides a measure of correlation between two variance terms. These models, known as random coefficient models, were evaluated for each periodontal outcome, yielding an overall mean trajectory across all sites, teeth, and subjects. All possible changes are summarized in the **Fig. 19.1**. Random coefficient

models were used to analyse longitudinal periodontal data consisting of repeated measures (level 1), sites (level 2), teeth (level 3), and subjects (level 4). Large negative and highly significant correlations between random linear and quadratic time coefficients indicated that subjects and teeth with greater-than-average linear change experienced decelerated variation. Conversely, subjects and teeth with less-than-average linear change experienced accelerated variation. Change therefore exhibited a dynamic regression to the mean at the tooth and subject levels. Since no equilibrium was attained throughout the study, changes were cyclical. When considered as a multilevel system, the ‘linear’ and ‘burst’ theories of periodontal disease progression are a manifestation of the same phenomenon: **Some sites improve while others progress, in a cyclical manner** (Gilthorpe et al. 2003).

19.6 Monitoring and Quality Control

A related design issue that needs to be addressed is the evaluation of examiner reliability. The validity and reproducibility of clinical diagnoses obtained in epidemiological investigations need to be documented whenever possible. Quality control measures are important because they directly impact on the estimates of disease and the magnitude of differences in disease levels between subpopulations due to induced bias and attenuation effects. It involves at least two major components: (1) training and calibration of examiners and (2) monitoring adherence to the protocol during the conduct of the study to assure that the data collected are valid and reliable (Kingman and Albandar 2002). The World Health Organization recommendation in *Oral Health Surveys; Basic Methods* (4th Edition, 1997) is that examiners taking part in epidemiological surveys should attend training and calibration sessions that should last for at least 4–5 days and should lead to intra- and inter-examiner agreement over scores in the range of 85–95% (Leroy et al. 2010). A more detailed discussion of quality control is given elsewhere (see chapter on **Measurement Reliability**).

As an example, background information on evaluations of data quality in terms of examiner reliability statistics (per cent agreements, kappa, and correlation coefficients) for the 1999–2002, 2003–2004, and 2005–2008 NHANES oral health component are provided.

19.6.1 The Oral Health Component of the 1999–2002 National Health and Nutrition Examination Survey (NHANES)

As Dye et al. (2007) summarized, since 1959, a series of health examination surveys of the US population have been conducted by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC). The initial surveys were named the national health examination surveys (NHES) and focused on adults (age 18–79 years) in 1959–1962, on children (age 6–11 years) in 1963–1965, and on adolescents (12–17 years) in 1966–1970. When an expanded dietary and nutritional assessment was added to the 1970–1972 examination cycle, the name of the survey was changed to the National Health and Nutrition Examination Survey (NHANES). Subsequently, a second National Health and Nutrition Examination Survey (NHANES II) was conducted in 1976–1980 and NHANES III followed during 1988–1994. During the period 1982–1984, a survey focusing on Hispanic populations in the US was conducted (HHANES). Oral health information has been collected during all these surveys, except for NHANES II (Dye et al. 2007 and references therein)

Beginning in 1999, NHANES was designed as a continuous survey with data released on a 2-year cycle to represent the civilian, non-institutionalized population of the US. The oral health component of the 1999–2002 National Health and Nutrition Examination Survey (NHANES) was a collaborative effort between the National Institute of Dental and Craniofacial Research (NIDCR), the National Center for Chronic Disease Prevention and Health Promotion, Division of Oral Health (NCCDPHP/DOH), and the National Center for Health Statistics (NCHS) (Dye et al. 2007).

Table 19.11 Comparison of sampling design characteristics for previous and current National Health and Nutrition Examination Surveys (NHANES) with oral health content (Dye et al. 2007) (Reprinted with permission from John Wiley & Sons, Inc.)

Characteristic	NHANES I	Hispanic HANES	NHANES III	NHANES 1999–2000	NHANES 2001–2002
Data collection year	1971–1974	1982–1984	1988–1994	1999–2000	2001–2002
Age of the target population	1–74 years	6 months to 74 years	2 months and older	From birth	From birth
Dental examination	Age 1–74 years	2–74 years	1 year and older	2 years or older	2 years or older
Number of survey examination locations	65	30	89	27	30
Eligible geographical areas for sample	48 lower continental states	Southwest for Mexican Americans; NY, NJ, CT for Puerto Ricans; Dade county Florida for Cubans	50 states	50 states + DC	50 states + DC
Groups targeted for expanded sampling	Persons 65 years or older; children 1–5 years of age; women aged 20–44 years	Persons 6 months to 19 years; and those aged 45–74 years	Persons 2 months < 6 years; non-Hispanic blacks; Mexican-Americans; those over 60 years	Persons 12–19 years; African-Americans; Mexican-Americans; those 60 years or older; and others ^{a,b}	Persons 12–19 years; African-Americans; Mexican-Americans; those 60 years or older; low-income whites; and others ^b
Total interviewed	28,043	13,689	33,994	9,965	11,039
Total MEC examined ^c	20,749	11,653	30,818	9,282	10,477
Total with a dental examination record	20,218	11,347	28,059	8,082	9,010

^aLow-income whites were over sampled only for NHANES 2000^bAll women who indicated that they were pregnant at time of screening were included in the sample^cMobile Examination Center (MEC)

Sample design characteristics for previous and 1999–2002 NHANES are shown in **Table 19.11**.

The specific objectives and related data uses for the 1999–2002 NHANES oral health data included (a) assessing the prevalence of major oral health diseases and conditions including dental caries, periodontal disease, dental trauma, dry mouth, and dental fluorosis; (b) assessing prevention and treatment efforts including the prevalence of dental sealants; (c) monitoring the oral health status of minority/underserved populations; (d) evaluating *Healthy People 2000 and 2010* objectives related

to oral health; and (e) supporting research activities as identified in the 2000 Surgeon General's Report on *Oral Health in America*. Details of the examination protocols and procedures for the 1999–2000 and 2001–2002 NHANES oral health component are available elsewhere (<http://www.cdc.gov/nchs/data/nhanes/oh-e.pdf>). The 1999–2000 and 2001–2002 oral health data sets are available at <http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm> (Dye et al. 2007).

Dental examiners for the 1999–2002 NHANES were dentists licenced in at least one US jurisdiction.

Other MEC personnel were trained as dental recorders. The quality of oral health data in this survey was controlled by an intense **training period** for dental examiners and recorders, calibration of dental examiners, and periodic monitoring and recalibration of dental examiners. In the training phase, the reference examiner (trainer) used lecture and slide presentations on each assessment to familiarize examiners with the study protocols and research criteria, including data recording and editing for that assessment. Infection control and emergency preparedness guidelines were reviewed as well. Demonstrations of examination technique and equipment use were also conducted (Dye et al. 2007).

In the **standardization phase**, the reference examiner and examiners in training examined the same set of volunteers. During this period, trainees were encouraged to ask questions regarding criteria while conducting the study protocols. A detailed discussion of observations was led by the reference examiner following each standardization round of examinations with the intent to systematically minimize differences in examination findings. The final phase of training included a preliminary calibration cycle with a follow-up calibration conducted in the field shortly after the dental examiner arrived for his/her first MEC stand. During calibration, examiners performed independent replicate examinations without discussion. Data from the calibration sessions were analysed to measure consistency between each examiner and the reference examiner. Training typically lasted for 40 h and was conducted in the Washington, DC, metropolitan area. The follow-up field calibration cycle was performed during a MEC's normal examination schedule and typically lasted for five to six 4-h examination sessions (20–24 h), depending on the number of study participants scheduled for those MEC sessions (Dye et al. 2007).

The reference examiner visited each dental examiner two to three times each year to observe field operations and to replicate randomly 20–25 dental examinations during each visit. The purpose of these periodic visits was to determine if the examiners were conducting the oral health examinations within the parameters of the study

protocols and if the standard for examination between the examiner and reference examiner had been maintained. Data from these replicate examinations were used to produce interrater reliability statistics. If an examiner's performance fell below an acceptable level, retraining was conducted on site. An examiner's performance was also assessed by monitoring results of second examinations. Approximately 5% of eligible study participants were asked to return to the MEC at a later time to have a number of assessments repeated, including the oral health exam. Data from the second exams were compared to the primary examination data to produce intrarater reliability statistics (Dye et al. 2007).

Weighted and non-weighted kappa statistics were produced using SAS software (version 8.02; SAS Institute Inc, Cary, NC, USA). For the Healthy People 2010 variable for prevalence of periodontal disease in the US (loss of attachment >4 mm at one or more sites), the strength of agreement between the primary dental examiners and the reference examiner from 1999 to 2002 varied from 0.58 to 0.77. For the overall subject-level inter- and intra-class correlation coefficients (ICC) measures, the interrater reliability ranged from 0.724 to 0.893 for loss of attachment and 0.550–0.866 for pocket depth (**Table 19.12**). Loss of attachment intra-class coefficients varied from 0.876 to 0.973. An analysis of potential examiner bias (systematic differences) regarding mean differences in loss of attachment/pocket depth between examiner and the reference examiner was performed. Mean differences in measurements across all periodontal sites ranged from 0.015 to 0.286 mm for attachment loss and from 0.013 to 0.524 mm for pocket depth (Dye et al. 2007).

19.6.2 The Oral Health Component of the 2003–2004 National Health and Nutrition Examination Survey (NHANES)

The 2003–2004 National Health and Nutrition Examination Survey (NHANES) was a collaborative effort involving 28 federal funding partners with the National Center for Health Statistics. The

Table 19.12 Dental examiner inter-classa and intra-classb correlation coefficients (ICC) for overall mean loss of attachment and mean pocket depth: NHANES, 1999–2002 (Dye et al. 2007) (Reprinted with permission from John Wiley & Sons, Inc.)

Characteristic	1999–2000 inter-rater				1999–2000 intra-rater				2001–2002 inter-rater				2001–2002 intra-rater			
	<i>n</i> ^c	Reference mean	Examiner mean	ICC	<i>n</i> ^c	First Ex mean	Second Ex mean	ICC	<i>n</i> ^c	Reference mean	Examiner mean	ICC	<i>n</i> ^d	First Ex Mean	Second Ex Mean	ICC
LOA mean ^d																
Examiner A	51	1.168	0.909	0.846	81	0.609	0.591	0.918	78	0.375	0.360	0.885	161	0.567	0.532	0.876
Examiner B	34	0.766	0.620	0.725	104	0.732	0.763	0.884	46	0.697	0.833	0.893	75	0.870	0.868	0.940
Examiner C									41	0.399	0.219	0.724	108	0.300	0.282	0.973
PD mean ^d																
Examiner A	51	1.626	1.613	0.866	81	1.498	1.547	0.749	78	0.785	0.811	0.744	161	1.063	1.035	0.827
Examiner B	34	1.256	1.009	0.550	104	1.158	1.178	0.756	46	1.072	1.337	0.600	75	1.241	1.256	0.865
Examiner C									41	0.884	0.656	0.637	108	0.700	0.680	0.827

^aPrimary examiner and reference examiner comparisons

^bFirst and second examination comparisons

^cNumber of paired observations

^dOverall mean is for the combined mesial and mid-facial periodontal sites

Table 19.13 Comparison of sampling design characteristics for previous and current National Health and Nutrition Examination Surveys (NHANES) with oral health content (Dye et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Characteristic	NHANES 1999–2000	NHANES 2001–2002	NHANES 2003–2004
Data collection year	1999–2000	2001–2002	2003–2004
Age of the target population	From birth	From birth	From birth
Dental exam age	2 years or older	2 years or older	2 years or older
Number of survey exam locations	27	30	30
Eligible geographical areas for sample	50 states + DC	50 states + DC	50 states + DC
Groups targeted for oversampling	Persons 12–19 years; non-Hispanic Blacks; Mexican-Americans; those 60 years or older; and others*	Persons 12–19 years; non-Hispanic Blacks; Mexican-Americans; those 60 years or older; low-income Whites; and others	Persons 12–19 years; non-Hispanic Blacks; Mexican-Americans; those 60 years or older; low-income Whites; and others
Total interviewed	9,965	11,039	10,122
Total mobile examination center examined	9,282	10,477	9,643
Total with an oral examination record	8,082	9,010	8,272

*Low-income Whites were oversampled only for the NHANES 2000.
DC District of Columbia

collaborators for the 2003–2004 NHANES oral health component included the National Institute of Dental and Craniofacial Research and the National Center for Chronic Disease Prevention and Health Promotion, Division of Oral Health. Oral health data are available on 8,272 persons aged 2 years or older. Sample design characteristics for the 2003–2004 NHANES are shown in **Table 19.13**. Additional information on the background and content of the NHANES can be found at <http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm> (Dye et al. 2008).

Valid questionnaire responses, code values, and criteria for assessment of dental conditions including tooth wear and functional contacts, as well as the 2003–2004 oral health data sets are available at <http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm>. Oral health derived variables in major reports, and SAS sample codes are available from the NIDCR/CDC Dental, Oral, and Craniofacial Data Resource Center at <http://drc.hhs.gov/Interrater> Evaluation (Dye et al. 2008).

Dental examiners undergo a comprehensive training and standardization period, including

periodic field reviews and repeat examinations, to ensure quality oral health data. For the 2003–2004 NHANES data collection cycle, second examinations conducted by the primary examiners were not performed. Consequently, the only reliability statistics available to assess data quality and examiner performance were derived from the repeat exams performed by the reference examiner on the study participants (interrater) (Dye et al. 2008).

Interrater data were collected on a convenience sample of NHANES study participants. During 2003–2004, the reference examiner evaluated the two primary dental examiners approximately four times each during regular mobile examination centre (MEC) operations. Site visits were evenly spaced and all sample participants scheduled for MEC examinations were eligible for the interrater reliability evaluation during these site visits. The reference examiner completed a repeat dental exam immediately following a standard oral health exam conducted by the primary dental examiner. The reference examiner determined which study participants to examine, and both

examiners were blinded to each other's observations. Given the constraints of appointing and examining study participants on the MEC, the primary dental examiners were aware of when interrater evaluations were to be performed. Although the calculated interrater agreements could have been affected by this knowledge, the interrater assessments were designed to minimize the potential for examiner bias (Dye et al. 2008).

The interrater reliability statistics produced for this report were per cent agreement, kappa statistics, and interclass correlation coefficients (ICCs). SAS software (version 8.02, SAS Institute Inc., Cary, NC) was used to calculate the weighted and non-weighted kappa statistics for a number of oral health conditions, including Healthy People 2010 derived oral health variables. A participant was identified as having periodontal disease if they had at least one periodontal site with 3 mm or more loss of attachment and 4 mm or more pocket depth at the same site. Periodontal status interrater reliability also was assessed by comparing ICCs generated from subject-level means (mm) for loss of attachment and pocket depth using measurements obtained from the three periodontal sites: mesio-facial, mid-facial, and disto-facial. Two primary dental examiners collected the majority of the 2003–2004 oral health data, compared with five examiners during 1999–2002. Examiner concordance for the NHANES 2003–2004 is consistent with the values reported for 1999–2002. The kappa scores for periodontal disease varied from 0.58 to 0.77 during 1999–2002 and from 0.64 to 0.72 in 2003–2004. The ICCs for mean pocket depth were similar as well. Coefficient values ranged from 0.55 to 0.87 during 1999–2002 and from 0.61 to 0.86 in 2003–2004. However, the ICC values for mean loss of attachment were slightly higher in 2003–2004 (0.86–0.93) compared with 1999–2002 (0.72–0.89) (Dye et al. 2008).

19.7 Self-Reported Periodontal Disease

The development, implementation, and evaluation of public health interventions for periodontal disease will require that the diseases be moni-

tored at several levels of the population. Current measures of periodontal disease are extremely resource intensive and cannot be used in several state-based surveillance systems. The existence and use of valid, low-cost, and low-resource self-reported measures of periodontal disease would be beneficial in a variety of ways. It would facilitate epidemiological studies of periodontal disease on a much larger scale than is feasible with the present clinical measures since much larger study populations could be reached by surveys rather than by clinical examination. Additionally, questions regarding periodontal disease could easily be added to ongoing studies to evaluate associations with other diseases and conditions. The use of self-report would allow for an easier and low-cost method of obtaining data for research and would support the creation of oral health programmes. Self-assessment can additionally serve as a motivational tool for good oral hygiene. Finally, self-reported measures would allow for surveillance of the periodontal condition of populations over time, in national, state, or regional surveillance programmes (Blicher et al. 2005).

19.7.1 Historical Perspectives on National Surveillance Activities for Periodontal Disease in United States

United States Public Health Service agencies began national surveillance activities for periodontal disease with the first National Health Examination Survey in 1960–1962, and this continued periodically through 2004 in the National Health and Nutrition Examination Survey (NHANES). For the assessment of periodontal disease in **NHANES I**, the Russell's periodontal index (PI) was used to promote the evaluation of trends in the epidemiology of periodontal disease between 1960 and 1962 and between 1971 and 1974. The PI was administered on all sampled participants aged 6–74 years. This was the only time that national efforts at periodontal disease surveillance included persons <13 years of age (Dye and Thornton-Evans 2007).

NHANES II was conducted by National Center for Health Statistics (NCHS) from 1976 to 1980 without an oral health examination. During this period, an expert nutrition panel recommended that NHANES conduct a study of health and nutrition status on a subpopulation of the United States following the conclusion of NHANES II using assessments similar to those employed in NHANES I and NHANES II. To address the panel recommendations and the inability to produce adequate prevalence estimates for Hispanics from prior national health examination surveys, NCHS conducted the **Hispanic Health and Nutrition Examination Survey (HHANES)** from 1982 to 1984. Three population subgroups of Hispanics were targeted (Mexicans, Cubans, and Puerto Ricans). An oral health examination similar to NHANES I was used on HHANES, with periodontal disease status assessed by the PI (Dye and Thornton-Evans 2007).

In 1988, all of the strengths and limitations of the NIDCR method for periodontal assessment were carried over to the next national health examination survey. **NHANES III** was conducted by NCHS from 1988 to 1994, and the oral health component of NHANES III was a collaborative effort between NIDCR and NCHS. The study design of NHANES III was unique compared to prior national health examination surveys. The study was designed to have two nationally representative study periods, or phases, of equal sample size and length. The descriptive findings from phase 1 (1988–1991) and later from the combined phases (1988–1994) represented the first reporting of probing depth, gingival recession, and attachment loss in the United States (Brown et al. 1996; Albandar et al. 1999). Although there were significant methodological differences used on NHANES III compared to prior national surveys for periodontal disease assessment that resulted in the inability to investigate trends in periodontal status over the past three decades, NHANES periodontal data have been used for two important national surveillance and health promotion-related activities: the **Surgeon General’s Report on Oral Health and Healthy People 2010** (Dye and Thornton-Evans 2007).

In 1999, NHANES was changed from a periodic survey to a continuous, annual survey using

a nationally representative sample for each year of data collection. The data are released in 2-year periods to protect confidentiality and to increase statistical reliability. The periodontal assessments used on NHANES beginning in 1999 were virtually unchanged from NHANES III and continued through 2004. From 1999 to 2000, the same two periodontal sites measured on NHANES III were assessed on NHANES 1999–2000. However, a third site, the distal-facial, was added in 2001 (Dye and Thornton-Evans 2007; Dye et al. 2007, 2011).

Clinically based periodontal examinations in NHANES ceased after the 2003–2004 data collection cycle. However, the established and suspected adverse effects of periodontal diseases led to the inclusion of a specific United States national objective in Healthy People 2010 (objective 21.5): ‘Reduce destructive periodontal disease in adults aged 35–44 years’. Although a national objective has been established, relatively few dental services are administered and delivered at the national level in the United States. Rather, most dental services are delivered through private dental offices or to a lesser extent, through state, county, or local dental public health programmes. Ideally, those public health programmes should be able to monitor disease levels, their distribution, and their trends within their jurisdiction to plan services and policies and evaluate their impact. However, surveillance for periodontal diseases is nearly non-existent at state, county, or local levels in the United States (Tomar 2007).

19.7.2 CDC Periodontal Disease Surveillance Project

CDC is the leading federal public health agency in the United States and is responsible for recommending surveillance methods that guide efforts to prevent and control disease and its risk factors. The DOH within CDC supports state- and community-based programmes and efforts to prevent oral disease by promoting science-based prevention strategies and monitoring oral health status and risk factors. In April of 2003, the CDC, in

collaboration with the American Academy of Periodontology (AAP), held a conference entitled 'Public Health Implications of Periodontal Infections in Adults'. Discussions at this conference emphasized the importance of surveillance of periodontal disease, as an important oral disease and as a potential risk factor for systemic diseases. Leading oral epidemiologists, biostatisticians, and periodontists attended this meeting. From these experts, a workgroup was convened and charged to examine and make recommendations for alternative valid population-based surveillance measures for periodontal disease that could be integrated into existing surveillance mechanisms such as the Behavioral Risk Factor Surveillance System (BRFSS). Thus, the primary objectives of this project were to determine whether self-reported measures can be valid measures for predicting the prevalence of periodontitis and if so, to identify and develop self-reported questions for use in the United States population (Eke and Genco 2007).

The 2007 *Journal of Periodontology* supplement contains a comprehensive summary of all work done by the CDC–AAP Periodontal Disease Surveillance Workgroup, including the accomplishments/activities presented at the 2006 IADR Symposium in Orlando, Florida. The consensus of the scientific evidence reported in this supplement suggests that multivariable modelling of self-reported measures is promising for predicting the prevalence of periodontitis and warrants further evaluation (Eke and Genco 2007).

19.7.3 Validity of Self-Reported Methods for Surveillance of Periodontal Disease

Several studies and a systematic review have been performed to validate self-reported measures for periodontal disease (Blicher et al. 2005; Glavind and Attström 1979; Nakashima et al. 1988; Tervonen and Knuuttila 1988; Unell et al. 1997; Gilbert and Nuttall 1999; Buhlin et al. 2002; Joshipura et al. 1996, 2002; Kallio et al. 1990, 1994; Kallio 1996; Nakashima et al. 1988, 1989; Tervonen and Knuuttila 1988; Pitiphat

et al. 2002; Genco et al. 2007; Gilbert and Litaker 2007; Dietrich et al. 2007; Taylor and Borgnakke 2007; Fisher et al. 2007). Although each of these papers analysed data sets from different sets of self-reported questions and from different represented populations and nationalities, increased validity using multivariable models was demonstrated consistently across their results (Eke and Genco 2007).

Of the many different self-reported periodontal items that were evaluated in these studies, self-reported tooth mobility was consistently found to be a highly significant predictor of both moderate and severe periodontitis using various different periodontal disease definitions (Dietrich et al. 2007, 2009; Gilbert and Litaker 2007; Slade 2007; Taylor and Borgnakke 2007). Self-reported tooth mobility by itself has a relatively low sensitivity but has consistently been shown to have a specificity of more than 90% for the diagnosis of moderate or severe periodontitis (Dietrich et al. 2007; Gilbert and Litaker 2007; Taylor and Borgnakke 2007; Dietrich et al. 2009). Dietrich et al. (2009) revealed that a simple prediction rule based on age, smoking, and self-reported tooth mobility may yield useful discrimination between subjects with and without a history of periodontitis and that such a prediction rule provides useful accuracy in a different population.

Blicher et al. (2005) reviewed the 16 studies that assessed the validity of self-reported periodontal and gingivitis measures against clinical gold standards. Seven of the studies included self-reported measures specific to gingivitis, four included measures only for periodontitis, and five included both gingivitis and periodontal measures. Three of the studies used a self-assessment method where they provided the patient with a detailed manual for performing a self-exam. The remaining 13 studies asked participants to self-report symptoms, presence of periodontal disease itself, or their recollection of a dental health professional diagnosing them or providing treatment for periodontal disease.

The results of the validation studies are presented in **Table 19.14** for periodontal disease and **Table 19.15** for gingivitis. Eight studies validated self-reported periodontal disease, and 13 studies

Table 19.14 Results from validation of self-reported periodontal disease (AQ) (Blicher et al. 2005) (Reprinted with permission from Sage Publications)

Self-report question	Clinical gold Standard	Results	Study reference
Periodontal disease			
1. Do you have ‘gum disease’?	Community Periodontal Index of Treatment Needs (CPITN)	38% clinically defined disease vs. 20% self-reported	Tervonen and Knuuttila (1988)
2. Thinks have gum disease	Any pockets >4 mm	SN=32%, SP=93%	Gilbert and Nuttall (1999)
	Any teeth with horizontal mobility >0.2 mm	SN=26%, SP=91%	
3. Do you have any periodontal/gum disease?	Above median % of sites with radiographic bone loss ≥20%	SN=19%, SP=86%	Pitiphat et al. (2002) (VADLS)
	>4 teeth with radiographic bone loss ≥40%	SN=18%, SP=84%	
Periodontal disease with bone loss			
4. Have you had periodontal disease with bone loss?	≥2 sites with radiographic bone loss >2 mm or complete loss of crestal lamina dura	PV ⁺ =76%, PV ⁻ =74%	Joshipura et al. (1996) (Dentists)
5. Have you had periodontal disease with bone loss?	Above median % of sites with radiographic bone loss ≥2 mm	PV ⁺ =80%, PV ⁻ =68%	Joshipura et al. (2002)
	Above median % of sites with radiographic bone loss ≥3 mm	PV ⁺ =72%, PV ⁻ =74%	(Non-dentists, HPFS)
6. Do you have periodontal disease or gum disease with bone loss?	Radiographic bone loss ≥2 mm	SN=39%, SP=100%, PV += 100%, PV -= 50%	Pitiphat et al. (2002) (HSDM)
Professional diagnosis of periodontal disease			
7. Has any dentist/hygienist told you that you have deep pockets?	# of pockets ≥4 mm (cut-off>8 pockets for ages 20–29 & 50–59, >10 pockets for ages 75–84)	^a SN=55%, ^a SP=90%, ^a PV ⁺ =77%, ^a PV ⁻ =75%	Buhlin et al. (2002)
8. Told by dentist/hygienist have gum disease	Any pockets >4 mm	SN=32%, SP=94%	Gilbert and Nuttall (1999)
	Any teeth with horizontal mobility >0.2 mm	SN=29%, SP=94%	
9. Have you ever been told by a dentist that you have periodontal/ gum disease with bone loss?	Above median % of sites with radiographic bone loss >20%	SN=33%, SP=82%	Pitiphat et al. (2002) (VADLS)
	>4 teeth with radiographic bone loss >40%	SN=50%, SP=78%	
Periodontal disease, severity			
10. In general, your current perio-dontal bone loss can be classified as: (none, mild, moderate, severe, don’t know)	None = Radiographic bone loss ≤1 mm, mild ≥1 mm but within coronal third of root, moderate = within middle third of root, severe = beyond middle third of root	Exact agreement=60.0%, % with ± one category = 76.0%, Spearman correlation=0.56	Pitiphat et al. (2002) (HSDM)

(continued)

Table 19.14 (continued)

Self-report question	Clinical gold Standard	Results	Study reference
Tooth mobility			
11. Highest recorded tooth mobility score (self-assessed)	Highest recorded tooth mobility	<i>SN=92%, SP=53%</i> ^a	Glavind and Attström (1979)
12. Think teeth loose or wobbly	Any pockets >4 mm Any teeth with horizontal mobility >0.2 mm	<i>SN=29%, SP=92%</i> <i>SN=32%, SP=94%</i>	Gilbert and Nuttall (1999)
13. # of sextants exhibiting bleeding or mobility (self-assessed) (also in Table 19.15⇔)	Pocket depth (>1 pocket deeper than 5 mm)	<i>P<0.02</i>	Glavind and Attström (1979)
Migrating teeth			
14. Changing position of front teeth during past year	# of teeth with Community Periodontal Index (CPI)=4	<i>b=0.012, P<0.001</i>	Unell et al. 1997
15. Think teeth have moved position	Any pockets >4 mm Any teeth with horizontal mobility >0.2 mm	<i>SN=18%, SP=83%</i> <i>SN=39%, SP=93%</i>	Gilbert and Nuttall (1999)
Recession			
16. Think can see more of roots of teeth than in past	Any pockets >4 mm Any teeth with horizontal mobility >0.2 mm	<i>SN=54%, SP=76%</i> <i>SN=54%, SP=78%</i>	Gilbert and Nuttall (1999)
Periodontal treatment			
17. Have you ever been told that you need periodontal or gum treatment?	Above median % of sites with radiographic bone loss >20% >4 teeth with radiographic bone loss >40%	<i>SN=47%, SP=68%</i> <i>SN=65%, SP=64%</i>	Pitiphat et al. (2002) (VADLS)
18. Have you ever had any form of periodontal or gum treatment?	Above median % of sites with radiographic bone loss >20% >4 teeth with radiographic bone loss >40%	<i>SN=48%, SP=65%</i> <i>SN=53%, SP=60%</i>	Pitiphat et al. (2002) (VADLS)
19. Aware of currently being treated for gum disease	Any pockets >4 mm Any teeth with horizontal mobility >0.2 mm	<i>SN=17%, SP=100%</i> <i>SN=13%, SP=99%</i>	Gilbert and Nuttall (1999)
Periodontal surgery			
20. Have you ever had periodontal surgery?	Above median % of sites with radiographic bone loss ≥2 mm Above median % of sites with radiographic bone loss ≥3 mm	<i>PV⁺=77%, PV⁻=70%</i> <i>PV⁺=69%, PV⁻=75%</i>	Joshiyura et al. (2002) (Non-dentists, HPFS)

^aCalculations done by us from data provided in manuscript
 No highlight or outline box: adequate statistics not reported
 Bold fonts: adequate statistics reported, but valid results not shown
 Bold italic fonts: adequate statistics reported and good validity shown

validated self-reported gingivitis and are included in both Tables. Five of these studies validated both self-reported periodontal disease and gingivitis (Glavind and Attström 1979; Nakashima et al. 1988; Tervonen and Knuuttila 1988; Unell et al. 1997; Gilbert and Nuttall 1999; Buhlin et al. 2002).

Results showing good validity are indicated by grey highlighting in **Tables 19.14** and **19.15**. An outline box indicates that the statistics given by

the authors were adequate but the validity was not satisfactory, and the absence of highlighting or an outline box indicates that inadequate

Table 19.15 Results from validation of self-reported gingivitis (Blicher et al. 2005) (Reprinted with permission from Sage Publications)

Self-report question	Clinical gold standard	Results	Study reference
Gingivitis			
1. Do you have gum problems?	Gingival bleeding index % (GBI%)	Self-reported 'no problems' show GBI%=6.1%, 'problems now and then' show GBI%=10.1%, 'often problems' show GBI%=24.5%	Schwarz (1989)
2. Do you have gingivitis now or have you ever had it before? ('currently' vs. 'never', excluded 'earlier' and 'can't say')	Bleeding on probing % (BOP%) CPITN Periodontal bleeding index (PBI%) # CPITN=0 sextants	$r=0.19, P<0.001$ $r=0.12, P<0.01$ $r=0.18, P<0.001$ $P<0.001$	Kallio et al. (1994)
Professional diagnosis of gingivitis			
3. Have you ever been diagnosed with gingival swelling or gingival problems by a dentist?	Any pockets >4 mm	$P<0.01$	Nakashima et al. (1988)
4. Have you ever been diagnosed with gingival swelling or gingival problems by a dentist?	Pocket depth	$P<0.001$	Nakashima et al. (1989)
Bleeding from gums			
5. # of sextants exhibiting bleeding or mobility (self-assessed) (also in Table 19.14↔)	>1 site with pocket>5 mm	$P<0.02$	Glavind and Attström (1979)
6. Have you noticed bleeding from your gums?	CPITN	98% clinically defined disease vs. 40% self-reported disease	Tervonen and Knuuttila (1988)
7. Trouble with bleeding gums	# of teeth with CPI=4	$\beta=0.14, P<0.001$	Unell et al. (1997)
8. Gums have bled sometimes	>40% sites bleed	SN=88%, SP=24%	Gilbert and Nuttall (1999)
9. Gums have bled recently	>40% sites bleed	SN=35%, SP=88%	Gilbert and Nuttall (1999)
10. Do your gums usually bleed?	Presence of BOP	SN=42%, SP=76%, aPV+=53%, aPV-=67%	Buhlin et al. (2002)
Bleeding on toothbrushing (self-assessed)			
11. Do you notice bleeding from your gums during toothbrushing?	Any pockets >4 mm	$P<0.01$	Nakashima et al. (1988)
12. Do you notice bleeding from your gums during toothbrushing?	Pocket depth	Not significant	Nakashima et al. (1989)
13. Have you ever noticed bleeding from your gums during toothbrushing? ('within 4 weeks' vs. 'never', excluded 'earlier' and 'can't say')	BOP% CPITN PBI% # CPITN=0 sextants	$r=0.26, P<0.001$ $r=0.19, P<0.001$ $r=0.2, P<0.001$ $P<0.001$	Kallio et al. (1994)

(continued)

Table 19.15 (continued)

Self-report question	Clinical gold standard	Results	Study reference
14. Bleeding on toothbrushing (self-assessed)	BOP%	<i>r=0.47, 74% agreement, kappa=0.27, SN=24%, SP=72%</i>	Kallio (1996)
15. Experience of gingival bleeding on brushing	Löe and Silness gingival index (Taani and Alhaija 2003)	<i>P=0.01, r=0.501</i>	Taani and Alhaija (2003)
Bleeding on toothpicking (self-assessed)			
16. Bleeding on toothpicking (self-assessed)	# of toothpick-induced bleeding fields	<i>P=0.005</i>	Glavind and Attström (1979)
17. Bleeding on toothpicking (self-assessed)	BOP%	<i>r=0.50, 73% agreement, kappa=0.31, SN=32%, SP=69%</i>	Kallio (1996)
18. Bleeding from gums while using toothbrush/toothpick (self-assessed)	Presence of BOP	<i>self-reported bleeding more likely to show high GB (P<0.001)</i>	Kallio et al. (1990)
19. Bleeding from gums while using toothbrush & toothpick (self-assessed)	BOP%	<i>r=0.54, 75% agreement, kappa=0.33, SN=31%, SP=71%</i>	Kallio (1996)
Inflammation of gums			
20. Do you have gum inflammation?	CPITN	98% clinically defined disease vs. 16% self-reported	Tervonen and Knuuttila (1988)
21. Do you have gingival swelling?	Any pockets >4 mm	<i>P<0.01</i>	Nakashima et al. (1988)
22. Do you have gingival swelling?	Pocket depth	<i>P<0.001</i>	Nakashima et al. (1989)
23. Think gums look swollen or puffy	>40% sites bleed	<i>SN=12%, SP=93%</i>	Gilbert and Nuttall (1999)
24. Symptoms of gingival bleeding and/or swelling	Alveolar bone support ratio ^b	<i>β=-5.95, P<0.01</i>	Taguchi et al. (1999)

No highlight or outline box: adequate statistics not reported

Bold fonts : adequate statistics reported, but valid results not shown

Bold italic fonts : adequate statistics reported, and good validity shown

^aCalculations done by us from data provided in manuscript

^bAlveolar bone support ratio is defined as the ratio of supporting bone height to root length (Salonen et al. 1991)

statistics were given for validity to be determined, and thus we were unable to evaluate these results. As seen in **Table 19.14**, self-reported measures for periodontal disease were grouped into several categories: disease awareness/perception as defined by the study participant (periodontal disease or periodontal disease with bone loss), knowledge of professional diagnosis of periodontal disease, severity of periodontal disease, symptoms of periodontal disease (tooth mobility, recession), and treatment (periodontal treatment, periodontal surgery) (Blicher et al. 2005).

The review indicates that some measures showed promise, but results varied across popu-

lations and self-reported measures. One example of a good measure is, 'Has any dentist/hygienist told you that you have deep pockets?', which had a sensitivity of 55%, a specificity of 90%, positive predictive value of 77%, and negative predictive value of 75% against clinical pocket depth. Higher validity could be potentially obtained by the use of combinations of several self-reported questions and other predictors of periodontal disease (Blicher et al. 2005).

Australian National Survey of Adult Oral Health (ANSAOH) provided interim results regarding the validity of six periodontal screening questions in predicting the prevalence of clinically

assessed periodontitis among Australian adults. The interview included six questions to screen for periodontal disease and five demographic/health history questions that represented traditional risk indicators for the disease. Oral examinations were conducted by trained, calibrated dentists who measured periodontal recession and probing depth at mesio-buccal, mid-buccal, and disto-buccal sites on all erupted teeth. A computer algorithm determined three categories of periodontal case status: none/mild, moderate, or severe. Multivariable binary logistic regression models were constructed using six screening questions: (1) Do you think that you have gum disease? (2). Has a dental professional ever told you that you have lost bone around your teeth? (3) Have you ever had scaling, root planing, surgery, or other treatment for gum disease? (4) Have you ever had any teeth that have become loose by themselves without some injury (not baby teeth)? (5) How often during the last 7 days did you use mouthwash or any dental rinse product? and (6) How often during the last 7 days did you use dental floss, tape, or an interdental brush to clean between your teeth, other than just to remove food particles stuck between your teeth?; five traditional risk indicators; and all 11 variables (Slade 2007).

Based on clinical findings, 4.0% of subjects were classified with 'severe' periodontitis and 25.2% were classified with 'moderate' periodontitis. Four percent of subjects met the 'severe' case definition for periodontitis, 25.2% met the 'moderate' case definition, and the remaining 70.8% of subjects met the definition for 'none/mild' periodontal disease. Correlations between pairs of the first four screening questions were weak (correlation coefficients ranging from 0.18 to 0.31), whereas correlations between those questions and the last two screening questions were nearly zero (range, 0.02–0.09). Four of the six screening questions were statistically significant predictors of a clinical diagnosis of moderate/severe periodontitis when assessed in a multivariable binary logistic regression model. However, the model's predictive validity, as indexed using combined sensitivity plus specificity, was only marginal (1.12). Four of the

five traditional risk indicators were statistically significant predictors of mild/moderate periodontitis, and together, they had better predictive validity. When all 11 variables were used to predict moderate/severe periodontitis, the overall validity improved (sensitivity plus specificity = 1.39). Four of the six screening variables remained statistically significant in this full model (Slade 2007).

Recently, the overall performance of eight self-report oral health questions in predicting the population prevalence of clinically determined periodontitis was assessed in a convenience sample of United States adults and included responses to these questions in the three largest racial/ethnic groups in the United States. Data were collected from 456 subjects participating in a 2007 study conducted by the Centers for Disease Control and Prevention. Each subject answered eight predetermined oral health self-report questions obtained from in-person interviews and were given a full-mouth periodontal examination using the National Health and Nutrition Examination Survey protocol. Three oral health questions, that is, scaling and root planing, having loose teeth, and use of mouthwash, yielded a sensitivity of 60% and a specificity of 72% in predicting total periodontitis in this sample with 22.4% total periodontitis. Tooth loss and risk and demographic factors seem to have a strong influence in predicting the prevalence of total periodontitis. About 80% of total periodontitis cases in this study were moderate periodontitis, which is more prevalent than severe periodontitis in the general population and thus is influenced more by the risk and demographic characteristics of the study populations. Also, some of these questions may be less applicable to mild periodontitis because they are related more to the late effects of periodontitis (e.g., loose teeth and lost bone around teeth). Multivariable modelling incorporating self-report measures on gum disease, loose teeth, and tooth appearance alone were most useful in predicting the prevalence of severe periodontitis and improved with the addition of demographic and risk factor variables, yielding an ROC value of 0.93, sensitivity of 54.6%, and specificity of 98% at the observed 4.8% prevalence of disease. Scaling and root

planing treatments, loose teeth, and the use of mouthwash, combined with demographic and risk factor covariates, were moderately useful in predicting total periodontitis (Eke and Dye 2009).

The cognitive testing of eight self-report questions was performed as part of a broader effort to evaluate and validate the use of these questions in estimating the prevalence of periodontitis in the United States population (Miller et al. 2007). The original questions were as follows: (1) Do you think you have gum disease? (2) Has a dental professional ever told you that you have lost bone around your teeth? (3) Have you ever had scaling, root planing, surgery, or other treatment for gum disease? (4) Have you ever had any teeth that have become loose by themselves without some injury (not baby teeth)? (5) How often during the last 7 days did you use mouthwash or any dental rinse product? (6) How often during the last 7 days did you use dental floss, tape, or interdental brush to clean between your teeth, other than just to remove food particles stuck between your teeth? (7) How would you rate the health of your gums? (8) During the past 3 months, have you noticed that you have a tooth that does not look right?

The primary difficulty with the question was that some respondents did not know what types of characteristics to include (specifically those unrelated to gum disease), and although they eventually answered 'no', some wondered if they should report 'yes'. Additionally, a few respondents stated that they simply do not look at their teeth and answered 'do not know'. The authors recommended for survey data to produce valid and reliable prevalence measures, questions must be understandable to respondents as well as cognitively possible for respondents to answer. Based on the results of cognitive testing, the recommended questions should be designed to reduce overall response burden, particularly for respondents of a large-scale national survey who would have little understanding of the causes or symptoms of gum disease. Additionally, the questions should be designed to be interpreted consistently by respondents and to be interpreted as the survey writers intended (Miller et al. 2007).

19.8 Using Administrative Data for Epidemiological Research

Automated administrative databases are omnipresent in medical and dental health care, and today's computing power allows sophisticated multi-layered analyses of care delivered to large populations (Spangler et al. 2012).

Investigators (Leake and Werneck 2005) have claimed that administrative data offer advantages, namely, that they:

1. Are generally longitudinal (data often span 10 years or more), based on clinical records which can be linked with other administrative data sets and/or over time to create a comprehensive representation of program use and client outcomes
2. Are detailed and accurate measures of program status and outcomes
3. Provide detailed information on the characteristics of participants, the services they have received, and the actions they have taken, which cannot be obtained in any other way
4. Are extensive and relatively inexpensive source of information
5. Contain data about the same conditions or programs over a long period, making it easy to determine if recent changes are unusual
6. Are readily available for timely analysis, retrieving needed information easily, which is critical for policy analyses
7. Can be a major tool in evaluating the impact of policy analyses and conducting strategic planning
8. Allow analysis of information about different patients and treatments, helping to create new research ideas
9. Offer '...benefits in the field of Health Services Research through the ordering of such a vast amount of information across so broad a spectrum...'

Disadvantages that occur with the use of administrative databases for research include (Leake and Werneck 2005):

1. Lack of agreement as to how items are coded.
2. Problems with accuracy and precision of data.

3. Patients are often incompletely characterized; key clinical data on processes and outcomes may be missing.
4. Identities and characteristics of providers may be inconsistently recorded.
5. Outcomes of interest may not be captured, for example, symptom relief, quality of life, out-of-hospital events, and level of satisfaction.

It was suggested that the addition of diagnosis codes, as well as Code on Dental Procedures and Nomenclature (CDT) codes, to dental claims (as is done with medical claims) would probably substantially improve the specificity as well as potentially allowing tracking of periodontal severity. The advent of electronic dental records that capture practitioner-entered periodontal probing depths means that such data could be systematically added to administrative records. Similar additions of blood pressure and body mass index enhance the usefulness of medical administrative data. At the same time, it may be possible to add other important markers of periodontal disease, such as attachment loss and furcation, to better reflect the American Academy of Periodontology gold standard periodontal disease case definition. Other useful additions would be systematic markers of edentulism and pharmacotherapy. However, it appears unlikely that dental claim codes can be used to distinguish the severity of periodontitis, such as mild, moderate, or severe. If CDT procedure codes are used as a surrogate measure for periodontitis, alternative measures would need to be employed to definitively differentiate periodontal disease and treatment effects (Spangler et al. 2012).

19.9 Risk Factors of Periodontal Disease

All persons are not equally susceptible to periodontal diseases and do not respond equally to periodontal therapy (Ronderos and Ryder 2004). Evaluating and understanding determinants of disease is central to recognizing the impact of risk factors on disease. Discussion of risk factors in the context of periodontal disease epidemiology typically focuses on two key

elements—identifying groups at risk and describing the magnitude of the increased risk. In large cross-sectional studies, the increased risk or a determinant's effect on disease outcome is assessed by evaluating the magnitude of the association between those members of a subgroup with the presence of the 'risk' factor compared to those without the factor under study and the occurrence of the event. In this context, a risk factor is more like a periodontal 'risk group', in that we are interested in identifying which subgroup(s) of a larger population is/are more likely to experience an adverse periodontal event compared to other members of the group (Dye 2012).

Above and beyond the identification of risk factors for periodontal diseases, it is important to assess the periodontal health risk from the presence of these factors and the effect of an increase in their dose. There is a considerable variation in the significance of the different factors to the initiation, modulation, and/or prediction of the occurrence and progression of periodontal tissue destruction. In addition, risk assessment may potentially provide profound social and economical benefits with regard to disease prevention and control (Albandar 2002).

Among the main measures used to express health risk are absolute risk, relative risk, odds ratio, and attributable risk (Albandar 2002):

- *Absolute risk*: is the probability that an individual will develop the disease over a specified period of time (Fig. 19.2).
- *Relative risk*: is the comparison of the health risk between two populations. This measure is commonplace in prospective and cohort studies and is assessed as the ratio of the disease incidence in these populations. A higher relative risk in these types of studies is often used to suggest a stronger evidence for causation of the outcome measure. It should be noted that relative risk can be measured only when the outcome is dichotomous.
- *Odds ratio*: is an estimate of the relative risk and is the ratio of the probability of occurrence of the event (or developing the disease) to the probability of its nonoccurrence. As this measure uses a backward method of analysis to compute the risk (from outcome to exposure),

Fig. 19.2 Risk assessment parameters (Albandar 2002. Reprinted with permission from John Wiley & Sons, Inc.)

		OUTCOME		
		+	−	
EXPOSURE	+	a	b	a + b
	−	c	d	c + d
		a + c	b + d	

Risk of outcome in exposed (absolute risk of exposed)	$= \frac{a}{a + b}$	Relative risk (RR)	$= \frac{\frac{a}{a + b}}{\frac{c}{c + d}}$
Risk of outcome in nonexposed (absolute risk of nonexposed)	$= \frac{c}{c + d}$	Attributable risk (AR)	$= \frac{a}{a + b} - \frac{c}{c + d}$
		Exposure odds ratio (OR)	$= \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{ad}{bc}$

it is useful in studies using a backward research design, including retrospective and case–control studies. Odds ratios have been commonly used to express risk because they provide a fairly good estimate of the true relative risk of exposure in the target population, provided that the outcome is rare.

- **Attributable risk:** is also a comparison of the health risk between two populations. However, in contrast to the relative risk, the attributable risk is assessed as the difference in the incidence rates (risk ratios or probabilities) of occurrence of disease between exposed and non-exposed individuals (or populations).
- **Population-attributable risk:** (also called *attributable risk fraction* or *etiologic fraction*) is an additional measure sometimes used to express the impact of exposure on outcome (or disease occurrence) in the target population from which the study sample derives. It measures the proportion of all cases of outcome in the target population that are attributable to exposure or the proportion of cases of outcome that would disappear (change to normal) if exposure in the target population were eliminated.
- **Prevalence rate ratio:** In cross-sectional studies, the selection of groups is by exposure, and

the outcome is measured as prevalence rather than incidence. In these studies, the *prevalence rate ratio* is often used as an approximation of the relative risk and is assessed as the ratio of the disease prevalence in the two populations.

Sensitivity and specificity are conceptually different but statistically related, indices. These indices are useful in the evaluation of a diagnostic test of disease and can be derived from a standard 2×2 table (Fig. 19.3). *Sensitivity* is the proportion of correctly identified diseased persons or the proportion of diseased persons who have a positive test. *Specificity* is the proportion of correctly identified disease-free persons or the proportion of disease-free persons who have a negative test (Albandar 2002).

Periodontal destruction is caused by uncontrolled host responses to pathogenic plaque biofilms. Some factors that may modify the balance of normal oral microbiota and host interactions may increase the likelihood of the individual's developing periodontal diseases. Some of these risk factors in the causative chain for periodontal diseases include poor oral hygiene, tobacco smoking, socio-economic status, malnutrition, psychological conditions, drug use, and local conditions, such as poor dental restorations and anatomical

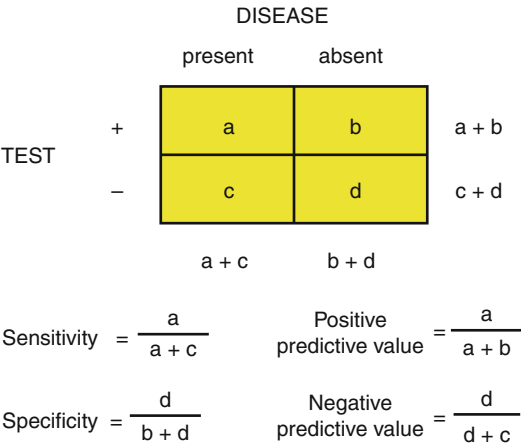


Fig. 19.3 Calculating the sensitivity and specificity of a diagnostic test (Albandar 2002. Reprinted with permission from John Wiley & Sons, Inc.)

Table 19.16 General and local established or potential risk factors of periodontal diseases

General/ systemic factors	Subject determinants
	Patient compliance and access to regular dental care
	Socio-economic status
	Tobacco smoking
	Uncontrolled diabetes mellitus
	History of periodontitis
	Neutrophil dysfunction and other acquired immunologic dysfunctions
	Genetic traits
	Human immunodeficiency virus (HIV) infection
	Medications (e.g., phenytoin, nifedipine, and cyclosporin)
	Stress
	Nutritional deficiency
	Hormones
	Obesity
	Osteoporosis
Local factors	Excessive alcohol consumption
	Poor oral hygiene
	Microbial factors
	Anatomic plaque-retentive factors
	Furcation involvements
	Enamel pearls and cervical enamel projections

Table 19.16 (continued)

	Root abnormalities (e.g., root grooves)
	Root proximity and open contacts
	Impacted third molars
	Overhanging restorations
	Pulpal involvement and root fracture
	External root resorption
	Trauma from occlusion
	Parafunctional habits
	Tooth mobility

Jin et al. (2011). Reprinted with permission from SAGE Publications

defects that make it difficult for individuals to perform oral hygiene. In addition, uncontrolled diabetes mellitus, obesity, untreated HIV infection, and genetic variables linked to hyper-inflammatory polymorphisms are risk factors (Table 19.16). Tobacco smoking is among the most important modifiable environmental risk factors for periodontitis. In this regard, periodontal diseases are complex multi-factorial conditions. Risk assessment and control are therefore fundamentally important in clinical practice (Jin et al. 2011).

References

Abd El-Aleem SA, Morales-Aza BM, Donaldson LF. Sensory neuropeptide mRNA up-regulation is bilateral in periodontitis in the rat: a possible neurogenic component to symmetrical periodontal disease. *Eur J Neurosci.* 2004;19:650-8.

Agerholm DM, Ashley FP. Clinical assessment of periodontitis in young adults-evaluation of probing depth and partial recording methods. *Community Dent Oral Epidemiol.* 1996;24:56-61.

Ainamo J, Ainamo A. Partial indices as indicators of the severity and prevalence of periodontal disease. *Int Dent J.* 1985;35:322-6.

Ainamo J, Barmes D, Beagrie G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) Community Periodontal Index of Treatment Needs (CPITN). *Int Dent J.* 1982;32: 281-291.

Albandar JM, Brunelle JA, Kingman A. Destructive periodontal disease in adults 30 years of age and older in the United States, 1988-94. *J Periodontol.* 1999;70: 13-29.

- Albandar JM, Streckfus CF, Adesanya MR, Winn DM. Cigar, pipe, and cigarette smoking as risk factors for periodontal disease and tooth loss. *J Periodontol*. 2000;71:1874–1881.
- Albandar JM. Global risk factors and risk indicators for periodontal diseases. *Periodontol* 2000. 2002;29:177–206.
- Albandar JM. Underestimation of periodontitis in NHANES surveys. *J Periodontol*. 2011;82:337–41.
- Aleksejunienė J, Holst D, Eriksen HM, Gjermo P. Psychosocial stress, lifestyle and periodontal health. *J Clin Periodontol*. 2002;29:326–35.
- Alpagot T, Wolff LF, Smith QT, Tran SD. Risk indicators for periodontal disease in a racially diverse urban population. *J Clin Periodontol*. 1996;23:982–8.
- Anagnou-Vareltzides A, Diamanti-Kipioti A, Afentoulidis N, Moraitaki-Tsami A, Lindhe J, Mitsis F, Papapanou PN. A clinical survey of periodontal conditions in Greece. *J Clin Periodontol*. 1996;23:758–63.
- Arbes Jr SJ, Agustsdottir H, Slade GD. Environmental tobacco smoke and periodontal disease in the United States. *Am J Public Health*. 2001;91:253–7.
- Armitage GC. Development of a classification system for periodontal diseases and conditions. *Ann Periodontol*. 1999;4:1–6.
- Armitage GC, Wu Y, Wang HY, Sorrell J, di Giovine FS, Duff GW. Low prevalence of a periodontitis-associated interleukin-1 composite genotype in individuals of Chinese heritage. *J Periodontol*. 2000;71:164–71.
- Al-Zahrani MS, Bissada NF, Borawski EA. Obesity and periodontal disease in young, middle-aged, and older adults. *J Periodontol*. 2003;74:610–5.
- Baelum V, Luan WM, Chen X, Fejerskov O. Predictors of destructive periodontal disease incidence and progression in adult and elderly Chinese. *Community Dent Oral Epidemiol*. 1997;25:265–72.
- Baelum V, Lopez R. Defining and classifying periodontitis: need for a paradigm shift? *Eur J Oral Sci*. 2003; 111:2–6.
- Baelum V, Fejerskov O, Manji F, Wanzala P. Influence of CPITN partial recordings on estimates of prevalence and severity of various periodontal conditions in adults. *Community Dent Oral Epidemiol*. 1993;21:354–9.
- Baelum V, López R. Defining a periodontitis case: analysis of a never-treated adult population. *J Clin Periodontol*. 2012;39:10–9.
- Baharin B, Palmer RM, Coward P, Wilson RF. Investigation of periodontal destruction patterns in smokers and non-smokers. *J Clin Periodontol*. 2006;33:485–90.
- Barratt H. 2009. <http://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/methods-of-sampling-population>. Accessed 24 Oct 2011.
- Beck JD, Löe H. Epidemiological principles in studying periodontal diseases. *Periodontol* 2000. 1993;2:34–45.
- Beck JD, Koch GG, Offenbacher S. Incidence of attachment loss over 3 years in older adults – new and progressing lesions. *Community Dent Oral Epidemiol*. 1995;23:291–6.
- Beck JD, Cusmano L, Green-Helms W, Koch GG, Offenbacher S. A 5-year study of attachment loss in community-dwelling older adults: incidence density. *J Periodontol Res*. 1997a;32:506–15.
- Beck JD, Sharp T, Koch GG, Offenbacher S. A study of attachment loss patterns in survivor teeth at 18 months, 36 months and 5 years in community-dwelling older adults. *J Periodontol Res*. 1997b;32:497–505.
- Beck JD, Caplan DJ, Preisser JS, Moss K. Reducing the bias of probing depth and attachment level estimates using random partial-mouth recording. *Community Dent Oral Epidemiol*. 2006;34:1–10.
- Blicher B, Joshupura K, Eke P. Validation of self-reported periodontal disease: a systematic review. *J Dent Res*. 2005;84:881–90.
- Borrell LN, Papapanou PN. Analytical epidemiology of periodontitis. *J Clin Periodontol*. 2005;32 Suppl 6:132–58.
- Borrell LN, Burt BA, Taylor GW. Prevalence and trends in periodontitis in the USA: from the NHANES III to the NHANES, 1988 to 2000. *J Dent Res*. 2005;84:924–30.
- Boström L, Linder LE, Bergström J. Influence of smoking on the outcome of periodontal surgery. A 5-year follow-up. *J Clin Periodontol*. 1998;25:194–201.
- Boström L, Linder LE, Bergström J. Smoking and cervical fluid levels of IL-6 and TNF-alpha in periodontal disease. *J Clin Periodontol*. 1999;26:352–7.
- Boström L, Linder LE, Bergström J. Smoking and GCF levels of IL-1beta and IL-1ra in periodontal disease. *J Clin Periodontol*. 2000;27:250–5.
- Boström L, Bergström J, Dahlén G, Linder LE. Smoking and subgingival microflora in periodontal disease. *J Clin Periodontol*. 2001;28:212–9.
- Breen HJ, Johnson NW, Rogers PA. Site-specific attachment level change detected by physical probing in untreated chronic adult periodontitis: review of studies 1982–1997. *J Periodontol*. 1999a;70:312–28.
- Breen HJ, Rogers P, Johnson NW, Slaney R. Option-4 algorithm for Florida pocket depth probe: reduction in the variance of site-specific probeable crevice depth measurements. *J Clin Periodontol*. 1999b;26:511–8.
- Breen HJ, Rogers PA, Slaney RE, Gillett IR, Johnson NW. Option-4 algorithm for third generation disc probe: agreement of selected site-specific relative attachment level measurements and detection of longitudinal site-specific attachment level change. *J Periodontol*. 1999c;70:159–70.
- Bridges RB, Anderson JW, Saxe SR, Gregory K, Bridges SR. Periodontal status of diabetic and non-diabetic men: effects of smoking, glycemic control, and socioeconomic factors. *J Periodontol*. 1996;67:1185–92.
- Brown LJ, Löe H. Prevalence, extent, severity and progression of periodontal disease. *Periodontol* 2000. 1993;2:57–71.
- Brown LJ, Oliver RC, Loe H. Periodontal diseases in the US in 1981: prevalence, severity, extent, and role in tooth mortality. *J Periodontol*. 1989;60:363–70.
- Brown LJ, Oliver RC, Loe H. Evaluating periodontal status of US employed adults. *J Am Dent Assoc*. 1990;121:226–32.
- Brown LJ, Brunelle JA, Kingman A. Periodontal status in the United States, 1988–91: prevalence, extent, and demographic variation. *J Dent Res*. 1996;75:672–83.

- Buhlin K, Gustafsson A, Andersson K, Hakansson J, Klinge B. Validity and limitations of self-reported periodontal health. *Community Dent Oral Epidemiol.* 2002;30:431–7.
- Burt B. Research, Science and Therapy Committee of the American Academy of Periodontology. Position paper: epidemiology of periodontal diseases. *J Periodontol.* 2005;76:1406–19.
- Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. *J Clin Periodontol.* 1986;13:500–505.
- Chen X, Wolff L, Aeppli D, Guo Z, Luan W, Baelum V, Fejeskov O. Cigarette smoking, salivary/gingival crevicular fluid cotinine and periodontal status. A 10-year longitudinal study. *J Clin Periodontol.* 2001;28:331–9.
- Chiappe V, Gomez M, Pedreira P, Galeano A, Grinfeld A, Viale J, Sznajder N. Longitudinal study of periodontal condition in students of the Dental School the University of Buenos Aires Argentina. *Acta Odontol Latinoam.* 1997;10:117–32.
- Costa FO, Guimarães AN, Cota LO, Pataro AL, Segundo TK, Cortelli SC, Costa JE. Impact of different periodontitis case definitions on periodontal research. *J Oral Sci.* 2009;51:199–206.
- Craig RG, Boylan R, Yip J, Bamgboye P, Koutsoukos J, Mijares D, Ferrer J, Imam M, Socransky SS, Haffajee AD. Prevalence and risk indicators for destructive periodontal diseases in 3 urban American minority populations. *J Clin Periodontol.* 2001;28:524–35.
- Cullinan MP, Westerman B, Hamlet SM, Palmer JE, Faddy MJ, Lang NP, Seymour GJ. A longitudinal study of interleukin-1 gene polymorphisms and periodontal disease in a general adult population. *J Clin Periodontol.* 2001;28:1137–44.
- Dietrich T, Stosch U, Dietrich D, Kaiser W, Bernimoulin JP, Joshipura K. Prediction of periodontal disease from multiple self-reported items in a German practice-based sample. *J Periodontol.* 2007;78:1421–8.
- Dietrich T, Kaiser W, Naumann M, Stosch U, Schwahn C, Biffar R, Dietrich D, Kocher T. Validation of a multivariate prediction rule for history of periodontitis in a separate population. *J Clin Periodontol.* 2009;36:493–7.
- Downer MC. The relative efficiencies of some periodontal partial recording selections. *J Periodontol Res.* 1972;7:334–40.
- Dowsett SA, Eckert GJ, Kowolik MJ. The applicability of half-mouth examination to periodontal disease assessment in untreated adult populations. *J Periodontol.* 2002;73:975–81.
- Dumitrescu AL, Abd-El-Aleem S, Morales-Aza B, Donaldson LF. A model of periodontitis in the rat: effect of lipopolysaccharide on bone resorption, osteoclast activity, and local peptidergic innervation. *J Clin Periodontol.* 2004;31:596–603.
- Dye BA. Global periodontal disease epidemiology. *Periodontol 2000.* 2012;58:10–25.
- Dye BA, Thornton-Evans G. A brief history of national surveillance efforts for periodontal disease in the United States. *J Periodontol.* 2007;78:1373–9.
- Dye BA, Barker LK, Selwitz RH, Lewis BG, Wu T, Fryar CD, Ostchega Y, Beltran ED, Ley E. Overview and quality assurance for the National Health and Nutrition Examination Survey (NHANES) oral health component, 1999–2002. *Community Dent Oral Epidemiol.* 2007;35:140–51.
- Dye BA, Nowjack-Raymer R, Barker LK, Nunn JH, Steele JG, Tan S, Lewis BG, Beltran-Aguilar ED. Overview and quality assurance for the oral health component of the National Health and Nutrition Examination Survey (NHANES), 2003–04. *J Public Health Dent.* 2008;68:218–26.
- Dye BA, Barker LK, Li X, Lewis BG, Beltrán-Aguilar ED. Overview and quality assurance for the oral health component of the National Health and Nutrition Examination Survey (NHANES), 2005–08. *J Public Health Dent.* 2011;71:54–61.
- Eaton KA, Duffy S, Griffiths GS, Gilthorpe MS, Johnson NW. The influence of partial and full-mouth recordings on estimates of prevalence and extent of lifetime cumulative attachment loss: a study in a population of young male military recruits. *J Periodontol.* 2001;72:140–5.
- Eke PI, Dye B. Assessment of self-report measures for predicting population prevalence of periodontitis. *J Periodontol.* 2009;80:1371–9.
- Eke PI, Genco RJ. CDC Periodontal Disease Surveillance Project: background, objectives, and progress report. *J Periodontol.* 2007;78:1366–71.
- Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA. Accuracy of NHANES periodontal examination protocols. *J Dent Res.* 2010;89:1208–13.
- Firatli E. The relationship between clinical periodontal status and insulin-dependent diabetes mellitus. Results after 5 years. *J Periodontol.* 1997;68:136–40.
- Fisher MA, Taylor GW, Shelton BJ, Debanne SM. Predictive values of self-reported periodontal need: National Health and Nutrition Examination Survey III. *J Periodontol.* 2007;78:1551–60.
- Fleiss JL, Park MH, Chilton NW, Alman JE, Feldman RS, Chauncey HH. Representativeness of the “Ramfjord teeth” for epidemiologic studies of gingivitis and periodontitis. *Community Dent Oral Epidemiol.* 1987;15:221–4.
- Gaengler P. Prevalence and distribution of gingivitis, periodontitis and missing teeth in adolescents and adults according to GPM/T index. *Community Dent Oral Epidemiol.* 1984;12:255–9.
- Galbraith GM, Hendley TM, Sanders JJ, Palesch Y, Pandey JP. Polymorphic cytokine genotypes as markers of disease severity in adult periodontitis. *J Clin Periodontol.* 1999;26:705–9.
- Genco RJ, Ho AW, Grossi SG, Dunford RG, Tedesco LA. Relationship of stress, distress and inadequate coping behaviors to periodontal disease. *J Periodontol.* 1999;70:711–23.
- Genco RJ, Falkner KL, Grossi S, Dunford R, Trevisan M. Validity of self-reported measures for surveillance of

- periodontal disease in two western New York population-based studies. *J Periodontol.* 2007;78:1439–54.
- Gilbert GH, Litaker MS. Validity of self-reported periodontal status in the Florida dental care study. *J Periodontol.* 2007;78:1429–38.
- Gilbert AD, Nuttall NM. Self-reporting of periodontal health status. *Br Dent J.* 1999;186:241–4.
- Gilthorpe MS, Zamzuri AT, Griffiths GS, Maddick IH, Eaton KA, Johnson NW. Unification of the “burst” and “linear” theories of periodontal disease progression: a multilevel manifestation of the same phenomenon. *J Dent Res.* 2003;82:200–5.
- Glavind L, Attström R. Periodontal self-examination. A motivational tool in periodontics. *J Clin Periodontol.* 1979;6:238–51.
- Gore EA, Sanders JJ, Pandey JP, Palesch Y, Galbraith GM. Interleukin-1beta+3953 allele 2: association with disease status in adult periodontitis. *J Clin Periodontol.* 1998;25:781–5.
- Grossi SG, Zambon JJ, Ho AW, Koch G, Dunford RG, Machtei EE, Norderyd OM, Genco RJ. Assessment of risk for periodontal disease. I. Risk indicators for attachment loss. *J Periodontol.* 1994;65:260–7.
- Grossi SG, Genco RJ, Machtei EE, Ho AW, Koch G, Dunford R, Zambon JJ, Hausmann E. Assessment of risk for periodontal disease. II. Risk indicators for alveolar bone loss. *J Periodontol.* 1995;66:23–9.
- Grossi SG, Skrepncinski FB, DeCaro T, Robertson DC, Ho AW, Dunford RG, Genco RJ. Treatment of periodontal disease in diabetics reduces glycated hemoglobin. *J Periodontol.* 1997;68:713–9.
- Haffajee AD, Socransky SS, Goodson JM. Comparison of different data analyses for detecting changes in attachment level. *J Clin Periodontol.* 1983;10:298–310.
- Heaton B, Dietrich T. Analytic epidemiology and periodontal diseases. *Periodontol 2000.* 2012;58:112–20.
- Holmgren CJ, Corbet EF, Lim LP. Periodontal conditions among the middle-aged and the elderly in Hong Kong. *Community Dent Oral Epidemiol.* 1994;22:396–402.
- Hugoson A, Jordan T. Frequency distribution of individuals aged 20–70 years according to severity of periodontal disease. *Community Dent Oral Epidemiol.* 1982;10:187–92.
- Hugoson A, Sjödin B, Norderyd O. Trends over 30 years, 1973–2003, in the prevalence and severity of periodontal disease. *J Clin Periodontol.* 2008;35:405–14.
- Hujoel PP, Leroux BG. Evaluating the burst hypothesis at a site-specific level using the lack-of-fit test. *J Periodontol.* 1998;69:357–62.
- Hunt RJ. The efficiency of half-mouth examinations in estimating the prevalence of periodontal disease. *J Dent Res.* 1987;66:1044–8.
- Hunt RJ, Fann SJ. Effect of examining half the teeth in a partial periodontal recording of older adults. *J Dent Res.* 1991;70:1380–5.
- Jeffcoat MK, Reddy MS. Progression of probing attachment loss in adult periodontitis. *J Periodontol.* 1991;62:185–9.
- Jin LJ, Armitage GC, Klinge B, Lang NP, Tonetti M, Williams RC. Global oral health inequalities: task group – periodontal disease. *Adv Dent Res.* 2011;23:221–6.
- Joshi KJ, Douglass CW, Garcia RI, Valachovic R, Willett WC. Validity of a self-reported periodontal disease measure. *J Public Health Dent.* 1996;56:205–12.
- Joshi KJ, Pitiphat W, Douglass CW. Validation of self-reported periodontal measures among health professionals. *J Public Health Dent.* 2002;62:115–21.
- Kaldahl WB, Johnson GK, Patil KD, Kalkwarf KL. Levels of cigarette consumption and response to periodontal therapy. *J Periodontol.* 1996;67:675–81.
- Kallio P. Self-assessed bleeding in monitoring gingival health among adolescents. *Community Dent Oral Epidemiol.* 1996;24:128–32.
- Kallio P, Ainamo J, Dusadeepan A. Self-assessment of gingival bleeding. *Int Dent J.* 1990;40:231–6.
- Kallio P, Nordblad A, Croucher R, Ainamo J. Self-reported gingivitis and bleeding gums among adolescents in Helsinki. *Community Dent Oral Epidemiol.* 1994;22(5 Pt 1):277–82.
- Kassab P, Colombier ML, Kaminski M, Lelong N, Sixou M, Nabet C. EPIPAP (EPIdeiological study on the relation between Periodontitis and Adverse Pregnancy outcomes) group. Impact of periodontitis definition in epidemiological research. Results from the EPIPAP study in postpartum women. *Eur J Oral Sci.* 2011;119:156–62.
- Kingman A, Albandar JM. Validity of partial recording of attachment loss in early-onset periodontitis. *J Periodontol Res.* 1997;76:229.
- Kingman A, Albandar JM. Methodological aspects of epidemiological studies of periodontal diseases. *Periodontol 2000.* 2002;29:11–30.
- Kingman A, Morrison E, Loe H, Smith J. Systematic errors in estimating prevalence and severity of periodontal disease. *J Periodontol.* 1988;59:707–13.
- Kingman A, Susin C, Albandar JM. Effect of partial recording protocols on severity estimates of periodontal disease. *J Clin Periodontol.* 2008;35:659–67.
- Laine ML, Farré MA, González G, van Dijk LJ, Ham AJ, Winkel EG, Crusius JB, Vandenbroucke JP, van Winkelhoff AJ, Peña AS. Polymorphisms of the interleukin-1 gene family, oral microbial pathogens, and smoking in adult periodontitis. *J Dent Res.* 2001;80:1695–9.
- Leake JL, Werneck RI. The use of administrative databases to assess oral health care. *J Public Health Dent.* 2005;65:21–35.
- Leroy R, Eaton KA, Savage A. Methodological issues in epidemiological studies of periodontitis-how can it be improved? *BMC Oral Health.* 2010;10:8.
- Locker D, Slade GD, Murray H. Epidemiology of periodontal disease among older adults: a review. *Periodontol 2000.* 1998;16:16–33.
- Løe H, Anerud A, Boysen H, Smith M. The natural history of periodontal disease in man. The rate of peri-

- odontal destruction before 40 years of age. *J Periodontol.* 1978;49:607–20.
- Lopez R, Baelum V. Classifying periodontitis among adolescents: implications for epidemiological research. *Community Dent Oral Epidemiol.* 2003;31:136–43.
- Laurell L, Romao C, Hugoson A. Longitudinal study on the distribution of proximal sites showing significant bone loss. *J Clin Periodontol.* 2003;30:346–52.
- Machtei EE, Christersson LA, Grossi SG, Dunford R, Zambon JJ, Genco RJ. Clinical criteria for the definition of established periodontitis. *J Periodontol.* 1992;63:206–14.
- Machtei EE, Hausmann E, Dunford R, Grossi S, Ho A, Davis G, Chandler J, Zambon J, Genco RJ. Longitudinal study of predictive factors for periodontal disease and tooth loss. *J Clin Periodontol.* 1999;26:374–80.
- Manau C, Echeverria A, Agueda A, Guerrero A, Echeverria JJ. Periodontal disease definition may determine the association between periodontitis and pregnancy outcomes. *J Clin Periodontol.* 2008;35:385–97.
- Martinez-Canut P, Lorca A, Magán R. Smoking and periodontal disease severity. *J Clin Periodontol.* 1995;22:743–9.
- McDevitt MJ, Wang HY, Knobelmann C, Newman MG, di Giovine FS, Timms J, Duff GW, Kornman KS. Interleukin-1 genetic association with periodontitis in clinical practice. *J Periodontol.* 2000;71:156–63.
- Meisel P, Siegemund A, Grimm R, Herrmann FH, John U, Schwahn C, Kocher T. The interleukin-1 polymorphism, smoking, and the risk of periodontal disease in the population-based SHIP study. *J Dent Res.* 2003;82:189–93.
- Meisel P, Kocher T. Definitions of periodontal disease in research: an alternative view. *J Clin Periodontol.* 2009;36:411–2.
- Miller A, Brunelle J, Carlos J, Brown L, Loe H, editors. Oral health of United States adults: the national survey of oral health in U.S. employed adults. Bethesda: National Institute of Dental Research, NIH Publication No. 87–2868; 1987.
- Miller K, Eke PI, Schoua-Glusberg A. Cognitive evaluation of self-report questions for surveillance of periodontitis. *J Periodontol.* 2007;78:1455–62.
- Mills WH, Thompson GW, Beagrie GS. Partial mount recording of plaque and periodontal pockets. *J Periodontol Res.* 1975;10:36–43.
- Mombelli A, Meier C. On the symmetry of periodontal disease. *J Clin Periodontol.* 2001;28:741–5.
- Mumghamba EG, Pitiphat W, Matee MI, Simon E, Merchant AT. The usefulness of using Ramfjord teeth in predicting periodontal status of a Tanzanian adult population. *J Clin Periodontol.* 2004;31:16–8.
- Nakashima K, Maeda S, Shimoyama M, Karami K, Shimojima T, Watanabe Y, et al. Epidemiological research of periodontal disease from questionnaire and pocket examination for junior and senior high school students in Kawagoe. *Nippon Shishubyo Gakkai Kaishi.* 1988;30:935–46.
- Nakashima K, Kurihara C, Kawanaga T, Kurihashi Y, Ohsawa K, Onodera O. Research into actual conditions and preventive care in periodontal disease. Relationship between questionnaire results and periodontal disease in youth. *Nippon Shishubyo Gakkai Kaishi.* 1989;31:1220–41.
- Norderyd O, Hugoson A. Risk of severe periodontal disease in a Swedish adult population. A cross-sectional study. *J Clin Periodontol.* 1998;25:1022–8.
- Norderyd O, Hugoson A, Grusovin G. Risk of severe periodontal disease in a Swedish adult population. A longitudinal study. *J Clin Periodontol.* 1999;26:608–15.
- Ogawa H, Yoshihara A, Hirotomi T, Ando Y, Miyazaki H. Risk factors for periodontal disease progression among elderly people. *J Clin Periodontol.* 2002;29(7):592–7.
- Owens JD, Dowsett SA, Eckert GJ, Zero DT, Kowolik MJ. Partial-mouth assessment of periodontal disease in an adult population of the United States. *J Periodontol.* 2003;74:1206–13.
- Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. *J Periodontol.* 2007;78(7 Suppl):1387–99.
- Paidi S, Pack AR, Thomson WM. An example of measurement and reporting of periodontal loss of attachment (LOA) in epidemiological studies: smoking and periodontal tissue destruction. *N Z Dent J.* 1999;95:118–23.
- Papapanou PN, Wennstrom JL, Johnsson T. Extent and severity of periodontal destruction based on partial clinical assessments. *Community Dent Oral Epidemiol.* 1993;21:181–4.
- Papapanou PN. Periodontal diseases: epidemiology. *Ann Periodontol* 1996;1:1–36.
- Papapanou PN, Baelum V, Luan WM, Madianos PN, Chen X, Fejerskov O, Dahlén G. Subgingival microbiota in adult Chinese: prevalence and relation to periodontal disease progression. *J Periodontol.* 1997;68:651–66.
- Papapanou PN, Neiderud AM, Sandros J, Dahlén G. Interleukin-1 gene polymorphism and periodontal status. A case-control study. *J Clin Periodontol.* 2001;28:389–96.
- Papapanou PN, Teanpaisan R, Obiechina NS, Pithpornchaiyakul W, Pongpaisal S, Pisuthanakan S, Baelum V, Fejerskov O, Dahlén G. Periodontal microbiota and clinical periodontal status in a rural sample in southern Thailand. *Eur J Oral Sci.* 2002;110:345–52.
- Parkhill JM, Hennig BJ, Chapple IL, Heasman PA, Taylor JJ. Association of interleukin-1 gene polymorphisms with early-onset periodontitis. *J Clin Periodontol.* 2000;27:682–9.
- Peng TK, Yao JH, Shih KS, Dong YJ, Chen CK, Pai L. Assessment of periodontal disease in an adult population survey in Taipei city using CPITN and GPM/T indices. *Chung-Hua Ya i Hsueh Hui Tsa Chih [Chin Dent J].* 1990;9:67–74.

- Peres MA, Peres KG, Cascaes AM, Correa MB, Demarco FF, Hallal PC, Horta BL, Gigante DB, Menezes AB. Validity of Partial Protocols to Assess the Prevalence of Periodontal Outcomes and Associated Socio Demographic and Behavior Factors in Adolescents and Young Adults. *J Periodontol*. 2011 Aug 22
- Pitiphat W, Garcia RI, Douglass CW, Joshipura KJ. Validation of self-reported oral health measures. *J Public Health Dent*. 2002;62:122–8.
- Preshaw PM. Definitions of periodontal disease in research. *J Clin Periodontol*. 2009;36:1–2.
- Querna JC, Rossmann JA, Kerns DG. Prevalence of periodontal disease in an active duty military population as indicated by an experimental periodontal index. *Mil Med*. 1994;59:233–6.
- Ramfjord SP. Indices for prevalence and incidence of periodontal disease. *J Periodontol*. 1959;30:51–59.
- Rodrigues DC, Taba MJ, Novaes AB, Souza SL, Grisi MF. Effect of non-surgical periodontal therapy on glycemic control in patients with type 2 diabetes mellitus. *J Periodontol*. 2003;74:1361–7.
- Ronderos M, Ryder MI. Risk assessment in clinical practice. *Periodontol* 2000. 2004;34:120–35.
- Salonen LW, Frithiof L, Wouters FR, Helldén LB. Marginal alveolar bone height in an adult Swedish population. A radiographic cross-sectional epidemiologic study. *J Clin Periodontol*. 1991;18:223–32.
- Savage A, Eaton KA, Moles DR, Needleman I. A systematic review of definitions of periodontitis and methods that have been used to identify this disease. *J Clin Periodontol*. 2009;36:458–67.
- Schwarz E. Dental caries, visible plaque, and gingival bleeding in young adult Danes in alternative dental programs. *Acta Odontol Scand*. 1989;47:149–57.
- Silness J, Røynstrand T. Partial mouth recording of plaque, gingivitis and probing depth in adolescents. *J Clin Periodontol*. 1988;15:189–92.
- Slade GD. Interim analysis of validity of periodontitis screening questions in the Australian population. *J Periodontol*. 2007;78:1463–70.
- Slade GD, Spencer AJ, Gorkic E, Andrews G. Oral health status and treatment needs of non-institutionalized persons aged 60+ in Adelaide, South Australia. *Aust Dent J*. 1993;38:373–80.
- Socransky SS, Haffajee AD, Goodson JM, Lindhe J. New concepts of destructive periodontal disease. *J Clin Periodontol*. 1984;11:21–32.
- Spangler L, Chaudhari M, Barlow WE, Newton KM, Inge R, Hujoel P, Genco RJ, Reid RJ. Using administrative data for epidemiological research: case study to identify persons with periodontitis. *Periodontol* 2000. 2012;58:143–52.
- Stewart JE, Wager KA, Friedlander AH, Zadeh HH. The effect of periodontal treatment on glycemic control in patients with type 2 diabetes mellitus. *J Clin Periodontol*. 2001;28:306–10.
- Susin C, Oppermann RV, Haugejorden O, Albandar JM. Periodontal attachment loss attributable to cigarette smoking in an urban Brazilian population. *J Clin Periodontol*. 2004;31:951–8.
- Susin C, Kingman A, Albandar JM. Effect of partial recording protocols on estimates of prevalence of periodontal disease. *J Periodontol*. 2005;76:262–7.
- Taani DQ, Alhaija ES. Self-assessed bleeding as an indicator of gingival health among 12–14-year-old children. *J Oral Rehabil*. 2003;30:78–81.
- Taguchi A, Suei Y, Ohtsuka M, Otani K, Tanimoto K, Hollender LG. Relationship between bone mineral density and tooth loss in elderly Japanese women. *Dentomaxillofac Radiol*. 1999;28:219–23.
- Taylor GW, Burt BA, Becker MP, Genco RJ, Shlossman M. Glycemic control and alveolar bone loss progression in type 2 diabetes. *Ann Periodontol*. 1998a;3:30–9.
- Taylor GW, Burt BA, Becker MP, Genco RJ, Shlossman M, Knowler WC, Pettitt DJ. Non-insulin dependent diabetes mellitus and alveolar bone loss progression over 2 years. *J Periodontol*. 1998b;69:76–83.
- Taylor GW, Borgnakke WS. Self-reported periodontal disease: validation in an epidemiological survey. *J Periodontol*. 2007;78(7 Suppl):1407–20.
- Tervonen T, Knuuttila M. Awareness of dental disorders and discrepancy between “objective” and “subjective” dental treatment needs. *Community Dent Oral Epidemiol*. 1988;16:345–8.
- Tervonen T, Karjalainen K. Periodontal disease related to diabetic status. A pilot study of the response to periodontal therapy in type 1 diabetes. *J Clin Periodontol*. 1997;24:505–10.
- Thomson WM, Williams SM. Partial- or full-mouth approaches to assessing the prevalence of and risk factors for periodontal disease in young adults. *J Periodontol*. 2002;73:1010–4.
- Thomson WM, Williams SM. Partial- or full-mouth approaches to assessing the prevalence of and risk factors for periodontal disease in young adults. *J Periodontol*. 2002;73:1010–4.
- Timmerman MF, van der Weijden GA, Armand S, Abbas F, Winkel EG, Van Winkelhoff AJ, Van der Velden U. Untreated periodontal disease in Indonesian adolescents. Clinical and microbiological baseline data. *J Clin Periodontol*. 1998;25:215–24.
- Timmerman MF, Van der Weijden GA, Abbas F, Arief EM, Armand S, Winkel EG, Van Winkelhoff AJ, Van der Velden U. Untreated periodontal disease in Indonesian adolescents. Longitudinal clinical data and prospective clinical and microbiological risk assessment. *J Clin Periodontol*. 2000;27:932–42.
- Timmerman MF, Van der Weijden GA, Hart AAM, Abbas F, Winkel EG, Van der Velden U. How do data from deepest pocket per quadrant relate to full-mouth scores? Progression of untreated periodontal disease in young Indonesians. *J Clin Periodontol*. 2002;29:219–23.
- Tomar SL. Public health perspectives on surveillance for periodontal diseases. *J Periodontol*. 2007;78:1380–6.

- Tonetti MS, Claffey N. European Workshop in Periodontology group C. Advances in the progression of periodontitis and proposal of definitions of a periodontitis case and disease progression for use in risk factor research. Group C consensus report of the 5th European Workshop in Periodontology. *J Clin Periodontol*. 2005;32 Suppl 6:210–3.
- Unell L, Söderfeldt B, Halling A, Paulander J, Birkhed D. Oral disease, impairment, and illness: congruence between clinical and questionnaire findings. *Acta Odontol Scand*. 1997;55:127–32.
- US Department of Health and Human Services. Focus area 21 oral health. In: *Healthy people 2010*. 2nd ed. Washington, DC: US Government Printing Office; 2000, p. 21–43. <http://www.healthypeople.gov/Document/pdf/Volume2/21Oral.pdf>. Accessed 26 Oct 2011.
- van der Velden U. Diagnosis of periodontitis. *J Clin Periodontol*. 2000;27:960–1.
- van der Velden U. Purpose and problems of periodontal disease classification. *Periodontol 2000*. 2005;39:13–21.
- Vettore MV, Lamarca Gde A, Leão AT, Sheiham A, Leal Mdo C. Partial recording protocols for periodontal disease assessment in epidemiological surveys. *Cad Saude Publica*. 2007;23:33–42.
- Wang WJ, Liu CY, Liu DZ, Lee CJ. Survey of periodontal disease among workers in Tianjin using Ramfjord's Periodontal Disease Index (PDI). *Community Dent Oral Epidemiol*. 1987;15:98–9.

Confounding and effect modification are of increasing importance as periodontal research addresses the putative associations between periodontal disease and systemic diseases. Unfortunately, the importance of confounding and effect modification is not always appreciated and is often overlooked. They may also appear to be similar but they are actually very different. Our goals in controlling them are also different. ‘In epidemiologic analysis one tries to eliminate confounding, but one tries to detect and estimate effect-measure modification’. We also assess effect modification first, since confounding by a variable becomes largely irrelevant if there is also interaction involving that variable (Hyman 2006).

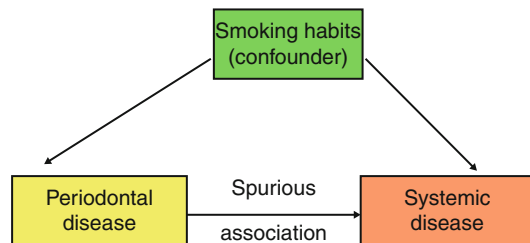


Fig. 20.1 An example of confounding

particular disease outcome. This implies that any apparent association between an exposure and a disease may be partly or completely accounted for by the confounder. Further, the absence of any apparent association between an exposure and a disease may result from the failure to account for the effects of a confounder. Thus, the distortion that results from confounding can lead to overestimation or underestimation of the true association, depending on the direction and magnitude of the relations that the confounder has with the exposure and disease variables. In very extreme cases, confounding can actually reverse the direction of an exposure effect; the latter is often referred to as *Simpson’s paradox* (Fitzmaurice 2003) (**Fig. 20.1**).

When examining the association between an exposure and a disease, it is important to question whether the apparent association may be due, at least in part, to the effects of one or more confounders. When considering the role of potential confounders, it is helpful to remember that,

20.1 Confounding

20.1.1 What Is a Confounder?

The term *confounder* has a couple of different meanings in everyday usage. For example, the *Oxford English Dictionary* provides two definitions of a confounder: ‘one who ruins, destroys, overthrows, spoils, discomfits’ and ‘one who causes confusion or disorder, who confuses distinctions’. It is the latter definition of the term that comes closest to its meaning as used in statistics. Statisticians and epidemiologists define a confounder to be a variable that distorts the true association between an exposure of interest and a

Table 20.1 Adjustment for tobacco smoking and the magnitude of periodontitis–systemic disease associations (Hujoel et al. 2002) (Reprinted with permission from John Wiley & Sons, Inc.)

Disease	Excellent ^a	Good ^b	Poor ^c
COPD^d	0.24 (0.90–1.72)	11.42 (1.16–1.72)	1.52 (1.21–1.91)
Lung cancer^e	0.58 (0.12–2.78)	1.48 (0.88–2.50)	1.94 (1.14–3.30)
Stroke^f	1.11 (0.79–1.57)	1.09 (0.82–1.45)	1.19 (0.84–1.73)
CHD^g	1.04 (0.82–1.32)	1.13 (0.95–1.34)	1.26 (1.02–1.56)

^aExcellent control for smoking refers to analyses limited to never-smokers (duration of smoking=0 years)

^bGood control for smoking was obtained by including smokers in the analyses (~50% of the population) and adjusting the analyses for the logarithm of smoking duration and the number of cigarettes per day

^cPoor control for smoking was obtained by limiting the analyses to smokers and not adjusting the analyses for smoking duration or dose

^dCOPD, chronic obstructive pulmonary disease (analyses were adjusted for age, age squared, race, poverty index, education, smoking duration and dose, and vitamins A and C)

^eLung cancer (analyses adjusted for the same variables as COPD analyses)

^fStroke (analyses were adjusted for same variables as CHD analyses)

^gCHD, coronary heart disease [analyses adjusted for age, age squared, gender, race (two indicator variables for African American and other), poverty index, marital state, education, and an interaction term for marital state and gender, diastolic blood pressure, systolic blood pressure, serum cholesterol, diabetes, log (height), log (weight), log (number of glasses per day), physical activity (indicator variable for heavy recreational or nonrecreational physical activity), and nervous breakdown and sampling design]. No adjustment for sampling weights was performed. The results differ by 1% from a previous report (11) (1.13 vs. 1.14) because 14 individuals missed in the previous analyses were included in the current analyses

for a variable to be considered a confounder of an association, it must satisfy three basic conditions:

1. The potential confounder must be associated with the disease outcome of interest.
2. The potential confounder must be associated with the exposure of interest.
3. The potential confounder must not be an ‘intermediate’ variable in the causal relation between the exposure and disease (Fitzmaurice 2003, MacKinnon et al. 2000).

20.1.2 Example of Confounders in Periodontal-Related Research

Age and sex are the most common confounding variables in health-related studies because these two variables are not only associated with most exposures we are interested in, such as diet, smoking habits, health beliefs, etc., but are also independent risk factors for many diseases (Needleman et al. 2005; Dietrich et al. 2008).

Hujoel et al. (2002) tested the hypothesis that the magnitude of the periodontitis–systemic disease associations is influenced by the prevalence of smokers and the degree of adjustment

for smoking. **Table 20.1.** shows a summary of the data on the periodontitis–systemic disease associations in three groups: smokers only, never smokers and smokers combined (but with adjustment for smoking history), and never smokers exclusively. The association between periodontitis and the systemic disease is reported using hazards ratios. A hazards ratio of 1.0 indicates a lack of association, a value above 1.0 indicates an elevated risk, and a value below 1.0 indicates a protective effect of periodontitis. If smoking is the driving factor behind the periodontitis–systemic disease associations, the magnitude of the associations should decrease as the control for smoking increases and the prevalence of smoking decreases.

Among past and current smokers, periodontitis significantly increased the risk for chronic obstructive pulmonary disease (COPD) [hazard ratio (HR) 1.52, 95% confidence interval 1.21–1.91]. When the analysis included past, current, and never smokers with adjustment for reported smoking dose and duration, the HR for COPD decreased by 7% (HR 1.42; 95% confidence interval 1.16–1.72). Finally, when the analysis was limited to never smokers, periodontitis was

associated with a small and insignificant increased risk for COPD (HR 1.24; 95% confidence interval 0.90–1.72) (Hujoel et al. 2002).

Among past and current smokers, periodontitis significantly increased the risk for lung cancer (HR 1.94; 95% confidence interval 1.14–3.30). When the analysis included never, past, or current smokers, with adjustment for smoking, the HR for lung cancer associated with periodontitis decreased by 49% (HR 1.48; 95% confidence interval 0.88–2.50). When the analysis was limited to never smokers, the opposite association was present. Periodontitis was associated with a decreased risk for lung cancer, not an increased risk (HR 0.58; 95% confidence interval 0.12–2.78) (Hujoel et al. 2002).

Among past and current smokers, periodontitis marginally increased the risk for stroke (HR 1.19; 95% confidence interval 0.84–1.87). When the analysis included never, past, or current smokers with adjustment for smoking, the HR for stroke associated with periodontitis decreased by 8% (HR 1.09; 95% confidence interval 0.82–1.45). Finally, when the analysis was limited to never smokers, periodontitis increased the HR for stroke by 11% (HR 1.11; 95% confidence interval 0.79–1.57) (Hujoel et al. 2002).

Among past and current smokers, periodontitis significantly increased the risk for CHD by 26% (HR 1.26; 95% confidence interval 1.02–1.56). Among past, current, and never smokers, the HR for CHD associated with periodontitis was 1.13 (95% confidence interval 0.95–1.34). Finally, when the analysis was limited to never smokers, the HR for CHD associated with periodontitis became insignificant and indistinguishable from a no-risk increase (HR 1.04; 95% confidence interval 0.82–1.32). The observed trends for these four systemic diseases are consistent with the hypothesis that imperfect adjustment for smoking history is inducing spurious associations between periodontitis and smoking-related diseases. Among never smokers, periodontitis was either not associated, or was weakly associated, with systemic diseases. None of the periodontitis–systemic disease associations were statistically significant among never smokers (Hujoel et al. 2002).

The recommendation to restrict periodontitis–systemic disease research to never smokers is a standard epidemiological research approach that is used whenever there is a strong confounding factor, the presence or absence of which is easily determined, but of which the magnitude cannot be well measured. Smoking definitely falls into this category (Hujoel et al. 2002).

20.1.3 Assessment of Tobacco Use in Periodontal Research

In clinical and epidemiological studies on the effects of smoking on periodontal disease, accurate measurement of smoking history is of central importance. Misclassification of smoking history will occur if a current smoker incorrectly reports that she has stopped smoking, a former smoker incorrectly recalls when he stopped smoking, or a current smoker underestimates the numbers of cigarettes he smokes per day, for example. Because smoking is such a common and strong risk factor for periodontitis, smoking is also a very important confounder in periodontal research, irrespective of whether periodontal status is the outcome or used as a predictor of systemic disease. Misclassification/measurement error with regard to smoking history in this scenario will result in **residual confounding** (Walter et al. 2012).

The resulting residual confounding may produce bias in the estimates of association between an exposure and outcome. In the setting of linear regression, one can show that the magnitude of the bias will be positively related to the amount of error involved in measurement of the confounder, the association between the confounder and the exposure, and the association between the confounder and the outcome, and inversely related to the variations of the confounder and of the outcome. Measurements of the confounder, tobacco exposure, are usually derived from participant self-reports, which have been reported to be suspect in terms of the numbers of cigarettes smoked per day. Tobacco exposure is a very important risk factor for periodontitis and is also strongly implicated in, if not the greatest risk factor for, many systemic

diseases currently linked with periodontitis. Rarely has there been a situation where a confounder has such a strong relation to both exposure and outcome. Combine this with the difficulties inherent in ascertaining accurate smoking information and one has a situation in which estimates of association may be heavily influenced by residual confounding (Spiekerman et al. 2003).

It has therefore been argued that **biomarkers** may be a more suitable measure of smoking exposure. However, biomarkers are of limited use in most clinical research settings. In contrast, biomarkers may be very useful in studies in which smoking status is the outcome of interest (i.e. cessation studies) or where verification of abstinence is required. Cotinine, thiocyanate, and carbon monoxide are the biochemical measures most frequently used to ‘validate’ smoking status and/or changes in status (Walter et al. 2012).

Modelling smoking history is not straightforward, as smoking is a multidimensional phenomenon with many characteristics such as intensity, duration, and time since cessation. Three different approaches can be distinguished (Walter et al. 2012):

1. Use of mutually exclusive categories of smoking history: The obvious advantages of this approach are its simplicity and the ease of interpretability of the resulting estimates.
2. Use of a composite measure of smoking intensity and duration (pack-years): This measure, used in an attempt to account for the intensity and duration of smoking exposure, is strictly cumulative and makes assumptions that are unlikely to hold in all situations. For instance, one pack-year of cigarettes smoked 20 years ago is assumed to have the same effect as the same amount smoked 1 year ago. Likewise, smoking five cigarettes per day over 20 years is assumed to have the same effect as 20 cigarettes per day smoked over 5 years.
3. Use of several indicator and/or continuous variables for different smoking characteristics: Alternatively, it is possible to include more than one characteristic (e.g. intensity, duration, time since cessation) as independent variables in a regression model. However, there are methodological limitations to this approach.

In aim to overcome these above-mentioned problems, a single, comprehensive smoking index that accounts for various characteristics of smoking as a function of its biological effect on disease risk over time was proposed. **The comprehensive smoking index (CSI)** accounts for changes in smoking patterns over time (Dietrich and Hoffmann 2004; Leffondré et al. 2006; Dietrich et al. 2007; Jimenez et al. 2009).

As Dietrich and Hoffmann (2004) assumed, the cigarettes smoked t years ago have an effect on chronic periodontitis which is proportional to $0.5^{t/\tau}$, where n is the number of cigarettes smoked *per* day and τ is a half-life parameter (in years). To allow for the effect of all years of active smoking, this term must be integrated over time.

Thus, the effect of smoking n cigarettes is proportional to

$$\text{CSI} = (1 - 0.5^{d/\tau}) \cdot 0.5^{c/\tau} \cdot n, \quad (20.1)$$

where d and c denote duration of smoking (in years) and time since cessation of smoking (in years), respectively. To account for changes in smoking over time (e.g. multiple relapses after cessation), one can split smoking history into k periods of constant exposure.

The CSI is then represented by

$$\text{CSI} = \sum_{i=1}^k (1 - 0.5^{d_i/\tau}) \cdot 0.5^{c_i/\tau} \cdot n_i \quad (20.2)$$

with

d_i – duration of i -th period (in years)

c_i – time since i -th period (in years)

n_i – number of cigarettes/day smoked in i -th period

The CSI is an increasing function of intensity (n) and duration (d), but a decreasing function of recency (c) (Dietrich and Hoffmann 2004).

Using NHANES III data from 12,623 subjects aged 20+ years, we compared the performance of the CSI with that of various conventional approaches using multiple logistic regression models of chronic periodontitis. The estimate of the smoking effect’s half-life was 1.5 years (95% CI, 0.5–2.5 years). Quantile plot of CSI by smoking category shows that CSI tends to increase with increasing smoking intensity (**Fig. 20.2**).

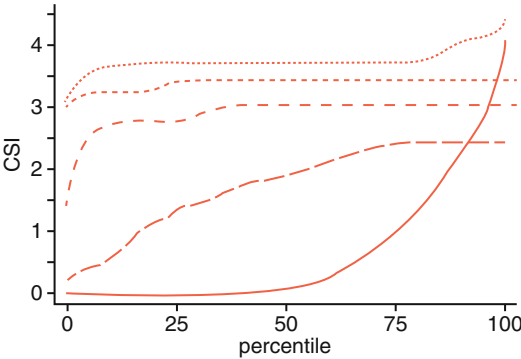


Fig. 20.2 Smoothed quantile plot of log-transformed CSI for former smokers (*solid line*) and different categories of current smokers [0–10 cig/day (*long dashes*), 11–20 cig/day (*medium dashes*), 21–30 cig/day (*short dashes*), and >30 cig/day (*dots*)] (Dietrich and Hoffmann 2004. Reprinted with permission from Sage Publications)

Furthermore, there is large variation within the category ‘former smokers’, due to variations in smoking intensity, duration, and recency. In addition, as the graph illustrates, due to variation in smoking duration, there is some overlap between current smoking categories, in particular, between lighter smokers (from 0 to 10 and 11 to 20 cigarettes per day). Depending on intensity, duration, and recency, the CSI can yield values that are considerably different from pack-years, a conventional composite smoking measure (Table 20.2). For example, a former smoker with 10 pack-years (20 cigarettes for 10 years) who quit 5 years ago can have a CSI of 1.96, while a former smoker with half the cumulative dose (10 cigarettes for 10 years) can have a CSI twice as high (3.93) if he quit only 2 years ago (Dietrich and Hoffmann 2004).

20.2 Controlling for Confounding

20.2.1 Control via Design

Randomization: By randomly allocating subjects to study groups, it is hoped that confounders are distributed equally between the groups (Needleman et al. 2005). The term random does not mean the same as haphazard but has a precise technical meaning. By random allocation we mean that each patient has a known chance, usually an

Table 20.2 Comparison of pack-years and CSI values for different values of smoking intensity, duration, and recency in current and former smokers (Dietrich and Hoffmann 2004) (Reprinted with permission from Sage Publications)

Intensity (cig/day)	Duration (years)	Recency (years)	Pack-years	CSI
Current smokers				
10	5	0	2.5	9.0
20	5	0	5	18.0
30	5	0	7.5	27.0
10	10	0	5	9.90
20	10	0	10	19.8
30	10	0	15	29.7
Former smokers				
10	10	2	5	3.93
20	10	2	10	7.86
10	10	5	5	0.98
20	10	5	10	1.96
10	10	10	5	0.10
20	10	10	10	0.19

equal chance, of being given each treatment, but the treatment to be given cannot be predicted. If there are two treatments, the simplest method of random allocation gives each patient an equal chance of getting either treatment; it is equivalent to tossing a coin. In practice, most people use either a table of random numbers or a random number generator on a computer. This is simple randomization. Possible modifications include block randomization, to ensure closely similar numbers of patients in each group, and stratified randomization, to keep the groups balanced for certain prognostic patient characteristics (Altman and Bland 1999). Randomization is usually the most effective way of minimizing the problem of confounding. If randomization is done properly, it has the advantage that it controls for both known and unknown confounders provided the sample size is sufficiently large (Needleman et al. 2005).

Restriction is the simplest way to control for a potential confounding factor. For example, if cigarette smoking is deemed a potential confounding factor in the relationship between oral contraceptives and cervical cancer, simply exclude smokers from the study. This neatly eliminates any effect that smoking might have on the relationship. Although valid, this approach is not popular for two practical reasons. First, this

exclusion limits the external validity of the findings: One can extrapolate the results only to women who do not smoke. Second, this approach inevitably reduces the sample size and thus the ability to detect effects of a given size should they exist (**power**) (Grimes and Schulz 2002).

Matching is another established approach to controlling for confounding. If smoking is deemed to be a potential confounding factor in a case–control study of oral contraception and cervical cancer, then for every case who smokes, a corresponding control who smokes is enrolled. Similarly, for every non-smoking case, a non-smoking control is selected. This manoeuvre makes the cases and controls homogeneous for smoking status. Although a valid approach to controlling confounding, matching can be cumbersome to do, slowing enrolment. In addition, an investigator cannot examine the effect of any confounding factor for which matching is done (Grimes and Schulz 2002).

20.2.2 Control via Analysis

Design-based methods are often infeasible or insufficient to produce exchangeability. Thus, there has been an enormous amount of work devoted to analytic adjustments for confounding. With a few exceptions, these methods are based on observed covariate distributions in the compared populations (Greenland and Morgenstern 2001).

Three methods for confounding assessment and adjustment for a non-mathematical readership were presented (Cleophas et al. 2006).

First method, **subclassification (stratification)**: The study population is divided into subclasses with the same subclass characteristic, then, treatment efficacy is assessed per subclass, and, finally, a weighted average is calculated (Cleophas et al. 2006).

Stratification has the advantage of creating subgroups that are more similar in terms of the baseline characteristics than the entire population, and this can result in less biased estimates of the intervention effect. However, stratification may reduce the power of the study to detect intervention effects because the total number of participants in each stratum will be reduced. Another limitation is

		Periodontitis	
		Present	Absent
Vitamin E Supplement	Yes	50	501
	No	65	384

Fig. 20.3 Hypothetical data on the association between vitamin E supplementation and the risk of periodontitis (Fitzmaurice 2004. Reprinted with permission from Elsevier)

that subgroups may not be balanced with respect to baseline risk factors, in which case the estimates of the intervention effect could still be biased. For this reason, stratification is often combined with regression techniques (Normand et al. 2005).

Figure 20.3 presents (hypothetical) data from a study that examined the association between vitamin E supplementation and the risk of occurrence of periodontal disease. Analysis of the data shown in **Fig. 20.3** indicates that the risk of periodontitis for those who used vitamin E supplements is approximately 0.09 (or 50 of 551) and that the corresponding risk for those who did not is approximately 0.15 (or 65 of 449). This apparent association can be expressed in terms of the odds ratio (OR): $OR = (50)(384)/(501)(65) = 0.59$. To address the concern of confounding, the data presented in **Fig. 20.3** are *stratified* according to whether or not a study participant is a smoker (**Fig. 20.4**). Focusing on the stratified data shown in **Fig. 20.4**, the odds of developing periodontitis among those who use vitamin E supplements, relative to those who do not, can be estimated separately for smokers and non-smokers. For smokers, the estimated OR is $OR (smokers) = (11)(200)/(40)(49) = 1.12$; for non-smokers the estimated OR is $OR (non-smokers) = (39)(184)/(461)(16) = 0.97$. These results indicate that there is little evidence of an association between vitamin E supplements and risk of periodontitis after controlling for the effects of smoking history (Fitzmaurice 2004).

Second method, **regression modelling**: In a multivariable regression model with treatment

Fig. 20.4 Hypothetical data on the association between vitamin E supplements and the risk of periodontitis, stratified by smoking history (Fitzmaurice 2004. Reprinted with permission from Elsevier)

		Periodontitis	
		Present	Absent
Smokers	Vitamin E Supplement		
	Yes	10	40
	No	50	200
		Present	Absent
Non-smokers	Vitamin E Supplement		
	Yes	40	460
	No	16	184

efficacy as independent and treatment modality as dependent variable, the covariates at risk of confounding are added as additional dependent variables to the model. An analysis adjusted for confounders is obtained by removing the covariates that are not statistically significant (Cleophas et al. 2006).

The main advantage of regression techniques is that they use data from all the participants. In addition, most researchers are familiar with these techniques and the analysis can be done using readily available software (Normand et al. 2005).

Third method, **propensity scores**: Each patient is assigned several odds ratios (ORs), which are his/her probability, based on his/her covariate value of receiving a particular treatment modality. A propensity score per patient is calculated by multiplying all of the statistically significant ORs. These propensity scores are, then, applied for confounding adjustment using either sub-classification or regression analysis (Cleophas et al. 2006).

A mathematical technique, termed the **Mantel–Haenszel procedure**, combines the various 2×2

tables (here, smokers and non-smokers) into a single summary estimate of the effect. The tables (or strata) are weighted inversely to their variance in deriving the summary statistic; stated alternatively, 2×2 tables with large numbers contribute more to the final statistic than do those with few patients. If the summary estimate of the treatment effect from the Mantel–Haenszel procedure differs substantially from the crude estimate from the overall data set, then confounding is deemed present (Grimes and Schulz 2002; Yu and Gastwirth 2008).

Suppose there are K levels of the stratifying variable (e.g., $K = 2$ for the data presented in **Figure 20.5**). The data can then be summarized in terms of K two-by-two contingency tables. The four internal cells of each table contain frequency counts of the number of individuals having a particular combination of the two variables. **Fig. 20.5** shows such a table, where the rows correspond to an exposure (e.g. vitamin E supplementation) and the columns correspond to disease status (e.g. presence or absence of periodontitis) (Fitzmaurice 2004). The Mantel–Haenszel formula for the pooled estimate of the OR from a

		Disease	
		Yes	No
Exposure	Yes	a_k	b_k
	No	c_k	d_k

Fig. 20.5 Stratum-specific two-by-two contingency table (Fitzmaurice 2004. Reprinted with permission from Elsevier)

series of K two-by-two contingency tables is given by

$$OR_{MH} = \frac{\sum_{k=1}^K (a_k d_k) / n_k}{\sum_{k=1}^K (c_k b_k) / n_k}$$

where n_k is the total number of observations in the k th table (i.e. $n_k = a_k + b_k + c_k + d_k$) and the summation in the formula is taken over the K levels of the stratification variable. For example, the MH pooled estimate of the OR based on the data from the two contingency tables shown in **Fig. 20.4** is

$$OR_{MH} = \frac{(11 \times 200) / 300 + (39 \times 184) / 700}{(40 \times 49) / 300 + (461 \times 16) / 700} = 1.03$$

As expected for the data shown in **Fig. 20.4**, the pooled estimate of the OR is approximately 1.0 and indicates that there is no evidence of association between vitamin E supplements and risk of periodontitis after controlling for the effects of smoking history (Fitzmaurice 2004).

20.3 Effect Modification

Effect modifying occurs whenever one or more variables and the exposure interact. For this reason, this phenomenon is also called ‘interaction’. Occurrence of interaction may be suggested by a different trend of the risk estimate within the confounder categories. In that case, due to the different association between risk and exposure in the

two analysed groups, a common risk estimate, adjusted for the effect of such a variable, cannot be calculated. Then interaction and confounding should always be considered as two very different phenomena (Parodi and Bottarelli 2005).

The occurrence of **interaction** may cause a reversal of the relative risk in the two categories of the effect modifier. In that case, the interaction is called ‘qualitative’, while when the risks estimated in the two exposure categories are both either above or below 1, the interaction is called ‘quantitative’, because the effect modifier influences the strength of the association, but not its direction. In some cases, it is possible to make a distinction between a ‘synergistic’ interaction, which occurs when the presence of the effect modifier enhances the impact of the exposure, and an ‘antagonist’ interaction, when the effect modifier reduces the effect of the exposure (Parodi and Bottarelli 2005).

In periodontology, one example of effect modification is that smokers experience less reduction in probing depth after non-surgical periodontal treatment compared with non-smokers (Labriola et al. 2005; Ylöstalo and Knuuttila 2006). Another example of effect modification is the association between cyclosporine medication and gingival overgrowth, where gingival overgrowth is modified by the presence of gingival inflammation (Pernu et al. 1992; Ylöstalo and Knuuttila 2006). A variation in the effect in different subgroups (effect modification) and a lack of comparability between exposed and unexposed subjects (confounding) can exist simultaneously (Ylöstalo and Knuuttila 2006).

Effect modification can be assessed using regression models or statistical tests such as a test of heterogeneity. However, it was suggested the use of stratified data as an interim tool in data analysis. In stratified data, stratum-specific estimates should be calculated first, and if effect modification is present, stratum-specific estimates should be reported since summary estimates do not convey information on the pattern of variation of stratum-specific estimates. In a situation where data are reasonably consistent, a singular estimate should be calculated either by summarizing stratum-specific estimates or by ignoring the stratification variable, depending on the situation, and the p -value for this should be

calculated (Rothman and Greenland 1998; Ylöstalo and Knuuttila 2006).

As an alternative to stratification, regression models with product term can be used to obtain stratum-specific estimates. For the sake of simplicity, many researchers do not report stratum-specific estimates if the variation in the estimates is small (Ylöstalo and Knuuttila 2006).

In addition to aetiological studies, effect modification affects the results of intervention studies. For example, the effect of treatment may be observed in groups of individual with certain characteristics, for example, among non-smokers or among study subjects with good compliance. If trials are constructed in such situations in a heterogeneous population in relation to these characteristics, it is possible that the beneficial effects are not be found or shown with sufficient precision. It may also be the case that the intervention may have a beneficial effect in one subgroup of patients, while it may have adverse effects in another subgroup (Ylöstalo and Knuuttila 2006).

References

- Altman DG, Bland JM. Statistics notes. Treatment allocation in controlled trials: why randomise? *BMJ*. 1999;318:1209.
- Cleophas TJ, Zwinderman AH, Cleophas TF, Cleophas EP. *Statistics applied to clinical trials*. 3rd ed. New York: Springer; 2006. p. 235–44.
- Dietrich T, Hoffmann K. A comprehensive index for the modeling of smoking history in periodontal research. *J Dent Res*. 2004;83:859–63.
- Dietrich T, Garcia RI, de Pablo P, Schulze PC, Hoffmann K. The effects of cigarette smoking on C-reactive protein concentrations in men and women and its modification by exogenous oral hormones in women. *Eur J Cardiovasc Prev Rehabil*. 2007;14:694–700.
- Dietrich T, Jimenez M, Krall Kaye EA, Vokonas PS, Garcia RI. Age-dependent associations between chronic periodontitis/edentulism and risk of coronary heart disease. *Circulation*. 2008;117:1668–74.
- Fitzmaurice G. Confused by confounding? *Nutrition*. 2003;19:189–91.
- Fitzmaurice G. Adjusting for confounding. *Nutrition*. 2004;20:594–6.
- Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health*. 2001;22:189–212.
- Grimes DA, Schulz KF. Clinical research in obstetrics and gynecology: a Baedeker for busy clinicians. *Obstet Gynecol Surv*. 2002;57(9 Suppl 3):S35–53.
- Hujoel PP, Drangsholt M, Spiekerman C, DeRouen TA. Periodontitis-systemic disease associations in the presence of smoking – causal or coincidental? *Periodontol* 2000. 2002;30:51–60.
- Hyman J. The importance of assessing confounding and effect modification in research involving periodontal disease and systemic diseases. *J Clin Periodontol*. 2006;33:102–3.
- Jimenez M, Krall EA, Garcia RI, Vokonas PS, Dietrich T. Periodontitis and incidence of cerebrovascular disease in men. *Ann Neurol*. 2009;66:505–12.
- Labriola A, Needleman I, Moles DR. Systematic review of the effect of smoking on nonsurgical periodontal therapy. *Periodontol* 2000. 2005;37:124–37.
- Leffondré K, Abrahamowicz M, Xiao Y, Siemiatycki J. Modelling smoking history using a comprehensive smoking index: application to lung cancer. *Stat Med*. 2006;25:4132–46.
- MacKinnon DP, Krull JL, Lockwood CM. Equivalence of the mediation, confounding and suppression effect. *Prev Sci*. 2000;1:173–81.
- Needleman I, Moles DR, Worthington H. Evidence-based periodontology, systematic reviews and research quality. *Periodontol* 2000. 2005;37:12–28.
- Normand SL, Sykora K, Li P, Mamdani M, Rochon PA, Anderson GM. Readers guide to critical appraisal of cohort studies: 3. Analytical strategies to reduce confounding. *BMJ*. 2005;330(7498):1021–3.
- Parodi S, Bottarelli E. Controlling for confounding in case-control studies. *Ann Fac Medic Vet Parma*. 2005; 25:19–46.
- Pernu HE, Pernu LM, Huttunen KR, Nieminen PA, Knuuttila ML. Gingival overgrowth among renal transplant recipients related to immunosuppressive medication and possible local background factors. *J Periodontol*. 1992;63:548–53.
- Rothman KJ, Greenland S. *Modern epidemiology*. 2nd ed. Philadelphia: Lippincott, Williams & Wilkins; 1998.
- Spiekerman CF, Hujoel PP, DeRouen TA. Bias induced by self-reported smoking on periodontitis-systemic disease associations. *J Dent Res*. 2003;82:345–9.
- Walter C, Kaye EK, Dietrich T. Active and passive smoking: assessment issues in periodontal research. *Periodontol* 2000. 2012;58:84–92.
- Ylöstalo PV, Knuuttila ML. Confounding and effect modification: possible explanation for variation in the results on the association between oral and systemic diseases. *J Clin Periodontol*. 2006;33:104–8.
- Yu B, Gastwirth JL. A method of assessing the sensitivity of the Cochran-Mantel-Haenszel test to an unobserved confounder. *Philos Transact A Math Phys Eng Sci*. 2008;366(1874):2377–88.

Animal models have been used to evaluate the pathogenesis of periodontal diseases and various periodontal treatment modalities. Human longitudinal studies of periodontal diseases pose many problems such as determining the level of disease activity, individuals at risk and susceptibility to disease progression. It is important to choose a laboratory animal model that has similar characteristics of human anatomy and periodontal disease. Features of periodontal diseases in humans and animals vary greatly depending upon which form of the disease is present and the stage of the development. Inflammatory destruction of the periodontium can be spontaneous or experimentally induced in most mammalian species (Weinberg and Bral 1999).

The selection of an experimental model is determined by research objectives, as well as laboratory constraints such as housing of large or nonstandard animals. The use of large animals with ethical and social issues such as monkeys and dogs should be reserved for last phase validation of new treatments prior to use in human clinical practice. In most cases, small animal models such as rats or hamsters will be sufficient to assess the role of bacteria, diet or other factors in periodontal inflammation at the histological level, providing sufficient statistical significance and preclinical relevance (Struillou et al. 2010).

21.1 Animal Models Used in Periodontal Research

Different animal models have been used in periodontal research, such as nonhuman primates, dogs, rabbits, ferrets, rats or hamsters, for periodontal diseases, biomaterials, dental implants or new regenerative strategies (Struillou et al. 2010). Variations between species include, but are not limited to, anatomy and dimensions of teeth and alveolar processes, amount and character of the gingiva, local physiological environment, animal behaviour and healing rate. In addition, species-specific feeding habits and other peculiar behaviour may interfere with healing of experimental sites. None of the species listed provide an anatomical and physiological environment equating that of the human (Table 21.1) (Wikesjö and Selvig 1999).

21.2 Selection of an Experimental Animal Model

The selection of an experimental model is determined by research objectives, as well as laboratory constraints such as housing of large or nonstandard animals. The use of large animals with ethical and social issues such as monkeys and dogs should be reserved for last phase validation

Table 21.1 High-level comparison of animal models of periodontal disease. The comparison is made based on general characteristics with major implications for the ease/difficulty in periodontal disease research and based on similarity with the disease in humans (Albuquerque et al. 2012) (Reprinted with permission from Elsevier)

Model	General characteristics				Similarity with the human disease					Well-characterized animal model
	Size	Cost	Handling	Availability	Dietary habits ^a	Anatomy ^a	Spontaneous periodontal disease ^a	Histopathology ^a	Oral microflora ^a	
Dog	++	-	++	-	++	+	++	+	+	++
Non-human primates	++	--	--	--	++	++	+	+	+	+
Rat; mouse	--	++	++	++	--	--	-	-	-	+
Hamster	--	++	++	++	--	--	-	-	-	--
Rabbit	+	+	++	++	--	-	-	-	-	--
Miniature pig	++	-	+	--	++	++	++	+	+	+
Ferret	-	+	+	+	--	-	+	+	+	-
Sheep	+	-	--	-	--	+	++	+	+	-

++ great advantage, + advantage, - disadvantage, -- great disadvantage

^aSimilarity with human characteristics

of new treatments prior to use in human clinical practice. In most cases, small animal models such as rats or hamsters will be sufficient to assess the role of bacteria, diet, or other factors in periodontal inflammation at the histological level, providing sufficient statistical significance and preclinical relevance (Struillou et al. 2010).

Effective and safe therapies for bone reconstructive therapies require preclinical evaluation to estimate their biologic potential, efficacy, and safety before clinical application and introduction. Candidate therapies should first be evaluated in well-characterized (critical-size) *screening models* for biologic potential and safety. Such models include ectopic and orthotopic rodent models. Promising therapies should then be evaluated for efficacy and clinical potential in discriminating large animal *critical-size defect models* in canines or nonhuman primates (Wikesjö et al. 2006). Critical-sized defects are defects that must not spontaneously regenerate following reconstructive surgery without adjunctive measures. In addition, it allows clinically relevant periodontal regeneration, induced or supported by implanted biologics, biomaterials, or devices over that in a surgical control. Once a candidate therapy has an established record of biologic potential and safety, and clinically relevant effect in a discriminating large animal critical-size defect model, successful candidate therapies may become subject to *clinical modelling* (Polimeni et al. 2006). Clinical modelling, that is, adaptation of clinical-type defects to experimental animals, is a common concept that has been applied to the study of induced, dehiscence, intrabony and class II or III furcation defects in canines and nonhuman primates. An advantage of clinical modelling is that defects of consistent morphology can be produced, facilitating histological processing and defect analysis. However, the peculiar anatomy and dimensions of teeth and alveolar processes in experimental animals and the possibility for extensive flap management and improved wound closure usually not encountered in the clinic detract value from clinical modelling. For example, study of experimental procedures in usually narrow, and shallow canine mandibular premolar

furcation defects with inherent short root trunks translate poorly to clinical furcation defects and thus have limited relevance. Nevertheless, canine and, in particular, nonhuman primate furcation defects have provided important data on the biological basis of periodontal wound healing (Wikesjö and Selvig 1999). Defects that may not necessarily be discriminating critical-size, but are recognized as difficult or desirable to manage, are produced in large animals to evaluate the clinical application and efficacy of a candidate therapy (Polimeni et al. 2006).

21.3 Experimental Periodontal Defects

Experimental defects commonly considered for reconstructive therapy have included defects caused by natural periodontal disease or defects induced by plaque-retaining crevicular ligatures (chronic defects). An alternative approach is the immediate use of surgically induced (acute) defects. **Natural periodontal disease defects** result from gradual, variable destruction and reduction of the periodontium and include infestation of bacterial endotoxins onto the root cementum (Haney et al. 1995; Wikesjö and Selvig 1999; Wikesjö et al. 1991).

Periodontal breakdown can be induced by **inoculation of microorganisms and their products** or by allowing plaque accumulation by placing a ligature around the cervical area of the molars (Duarte et al. 2010). **Root surfaces of ligature-induced or other chronic defects** are expected to harbour similar qualities with the natural periodontal disease defects. Briefly, a cotton, silk or nylon ligature is placed in the cervical area of the maxillary or mandibular rat molars to allow plaque accumulation, inducing changes in the periodontal tissues similar to those observed in human periodontitis, including rupture and apical migration of junctional epithelium, influx of inflammatory cells and periodontal ligament fibres and bone loss (Figs. 21.1 and 21.2). Most of the studies using this methodology keep the ligature in position from 15 to 60 days in order to induce periodontal destruction. At 42 days, the

ligature biofilm in rats presents various bacterial species that hybridize to probes of periodontal bacterial species commonly observed in human (Duarte et al. 2010).

Conventional root debridement, however, eliminates most, if not all, of the cementum in natural disease or chronic defects, resulting in root surface conditions similar to those of acute defects (Wikesjö and Selvig 1999; Wikesjö et al. 1991; Jones and O'Leary 1978). Thus, from the perspective of root surface characteristics, the rationale for selecting natural disease or chronic models appears diluted. Natural disease and chronic defects are characterized by compromised gingi-

val dimensions and character, which may provide some clinical relevance but represent a confounding factor in evaluation of the healing results (Wikesjö and Selvig 1999; Wikesjö et al. 1991).

The surgically created acute periodontal defect was preferable to a natural plaque-induced chronic periodontal defect because the surgical preparation of a defect can save time and allows standardization of defect size. In addition, meticulous root planing would detoxify the root surface and, consequently, would provide the same post-surgical healing to both acute and chronic defects (Jung et al. 2011). A desirable defect model should be standardized, reproducible, and easily prepared (Kim et al. 2004).

Several types of experimental defects have been utilized to evaluate specific aspects of periodontal wound healing, for example, the **periodontal fenestration defect** (Fig. 21.3 and 21.4) (Huang et al. 2005; Tal et al. 2005; Benatti et al. 2006; Corrêa et al. 2010; Sculean et al. 2008; Artzi et al. 2006; Zhao et al. 2004; Caplanis et al. 1998; King et al. 1997, 1988; Choi et al. 1993). This rat periodontal defect model has proven to be a predictable and reliable model in studying periodontal healing without oral bacterial contamination or ingrowth of gingival epithelium. Although this model may not represent a critical-size defect for periodontal regeneration, it serves as a reasonable screening model for using to examine the wound-healing kinetics. A very

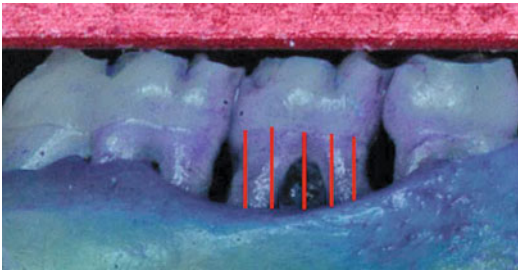


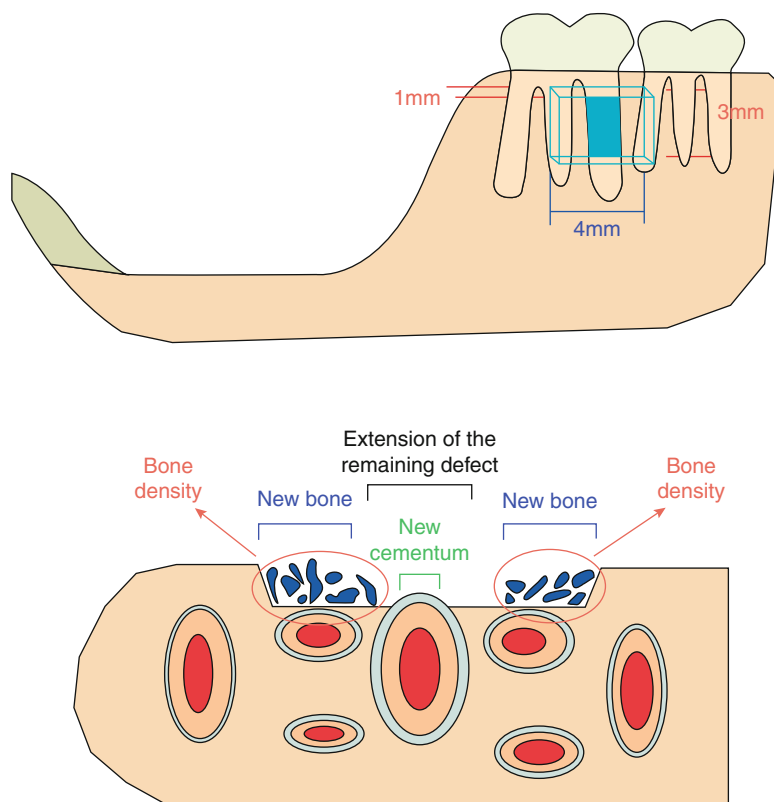
Fig. 21.1 Alveolar bone loss measurement defleshed jaws in a ligature-induced periodontitis rat model. Periodontal bone loss was defined as the distance between the cemento-enamel junction and the alveolar bone crest. Buccal and palatal measurements were made at five points and a mean of these values was considered as the bone loss (Lieberman et al. 2011. Reprinted with permission from Elsevier)



Fig. 21.2 Photomicrographs illustrating bone loss in the furcation area in a molar that received ligature (a) and the periodontal ligament area in an unligated molar (b). The

yellow line is delimitating the area of bone loss. The green line is delimitating the area of periodontal ligament (Duarte et al. 2010. Reprinted with permission from Elsevier)

Fig. 21.3 Diagram showing surgical creation of the periodontal defect and the evaluated parameters (the percentage of bone fill (*BF*), the density of newly formed bone, new cementum formation (*NC*) and extension of the remaining defect (*ERD*) histomorphometrically obtained by using an image analysis system) in a rat animal model (Benatti et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)



interesting finding of the present study, and of others, is the fact that despite the absence of ‘contamination’ with dental biofilm, no new cementum was formed along the root surface in any of the above-mentioned groups (Benatti et al. 2006). By using this model, cementum regeneration has been mostly seen if any regenerative technique or collagen carrier was applied (King et al. 1997; Zhao et al. 2004). This observation raises the point that cementum presents a very limited self-healing capacity, even in the absence of previous exposure to the dental biofilm (Benatti et al. 2006).

Other defect models have adopted **submerged wound closure** to study the regenerative capacity of the periodontium. It is essential with any experimental defect considered to portray healing events in the periodontium or elsewhere in the skeleton that it clearly satisfies criteria for a critical-size defect; that is, that the defect will not regenerate within its lifetime without adjunctive measures. Use of smaller defects may only detract

from understanding periodontal wound healing and regeneration (Wikesjö and Selvig 1999).

The critical-size supra-alveolar periodontal defect model was developed and characterized using the young adult, male, beagle dog. Following appropriate animal preparation, the alveolar bone was surgically reduced to a level of 5–6 mm from the cemento-enamel junction around the third and fourth mandibular premolar teeth and the exposed cementum removed. The first and second premolar teeth were extracted and the first molar amputated at the level of the reduced alveolar crest. The resulting circumferential, supra-alveolar periodontal defect was used for immediate reconstructive intervention (Fig. 21.5). This procedure results in a critical-size defect. Spontaneous alveolar bone and cementum regeneration does not exceed 25% of the defect height over a 4-week healing interval following mucogingival flap surgery. Extending the healing interval to 8 weeks does not result in further regeneration, nor does the mode of wound

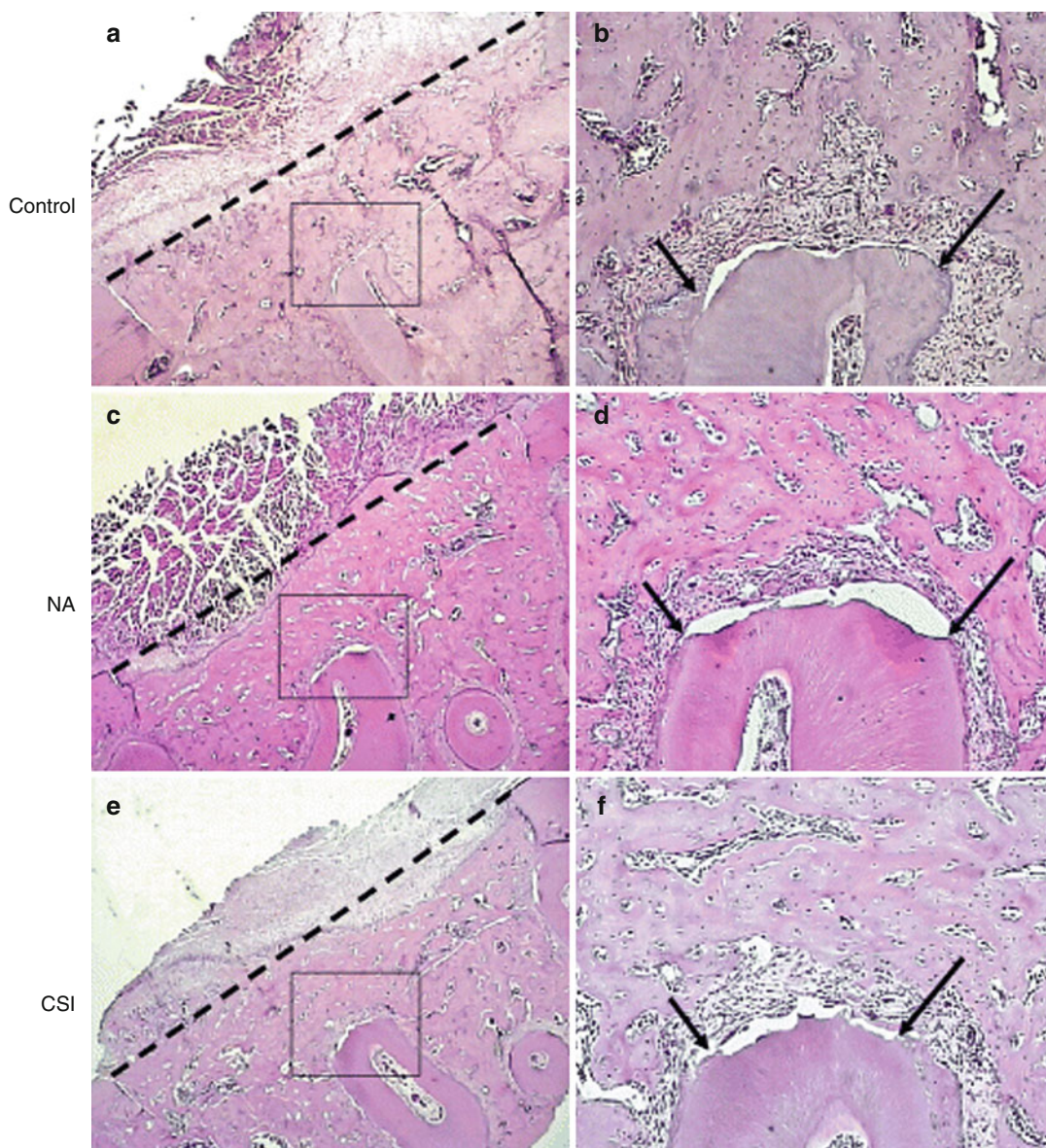


Fig. 21.4 Histological illustration of periodontal healing 21 days after the surgical creation of the fenestration-type defects (H&E staining). *Lines* in figures (a, c, e) correspond to the original extension of the defects, while *arrows* in figures (b, d, f) indicate the margins of the pre-existing cementum. Figures (a, b) represent group con-

trol, (c, d) represent group NA (nicotine administration), and (e, f) represent group CSI (cigarette smoke inhalation) (original magnification (a, c, e): 5 $\times$; (b, d, f): 20 $\times$) (Benatti et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

closure, that is, transgingival or submerged. The defect dimensions provide for clinically relevant regeneration of alveolar bone and cementum (Fig. 21.6) (Wikesjö and Selvig 1999). The effect of defect characteristics and space provision,

innate regenerative potential and healing dynamics has received further analysis using the critical-size supra-alveolar periodontal defect model (Fig. 21.7) (Polimeni et al. 2004a, b, c, d, 2005, 2009; Wikesjö et al. 2003a, b).

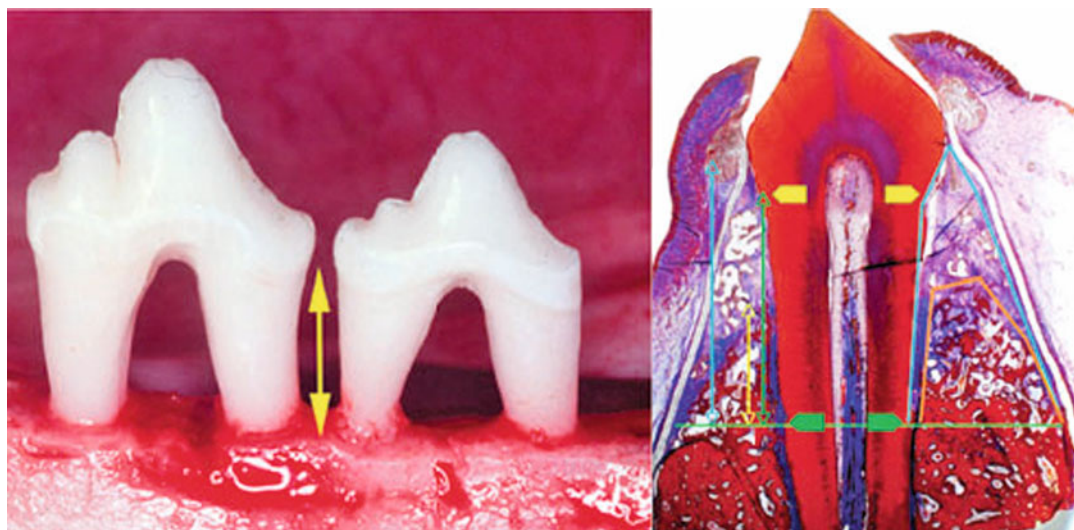


Fig. 21.5 The critical-size, supra-alveolar periodontal defect model. The alveolar bone and periodontal attachment, including the cementum, are surgically reduced circumferentially around the third and fourth mandibular premolar teeth to a level 5–6 mm from the cemento-enamel junction. The first molar is reduced to the level of the reduced alveolar bone and the first and second premolars are extracted. Experimental treatments are applied immediately upon defect induction. Wound closure for primary intention healing may be transgingival, leaving the tooth structure intact, or submerged following reduction of the clinical crowns. Examples of histometric parameters evaluated in the critical-size, supra-alveolar

periodontal defect model are shown: The *green line and arrowheads* represent the base of the surgically created defect and the *yellow arrowheads* represent the cemento-enamel junction. The defect height (*vertical green arrow*), bone regeneration height (*vertical yellow arrow*), defect area (*blue lines*) delineated by an expanded polytetrafluoroethylene (ePTFE) membrane in this example (membrane height: *vertical blue arrow*), and bone regeneration area (*orange lines*) are shown. The white irregular ‘ghost’ structures within the wound area and regenerated alveolar bone represent a bone biomaterial evaluated in this example (Polimeni et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

A method involving calculating the mean value of histometric assessments of several step-serial histologic sections obtained from the same site traditionally has been used for the histometric characterization of the site. The histometric analysis of the supra-alveolar periodontal defect model uses observations from buccal and lingual sites from three step-serial sections representing the central 0.2-mm aspect of the mesial and distal root for two mandibular premolar teeth (Koo et al. 2004a, b). The following histometric parameters were assessed separately for the buccal and the lingual sections of each site (Fig. 21.6) (Koo et al. 2004a, b):

1. Defect height: distance between the apical extension of root planing and the cemento-enamel junction
2. Defect area: area under the ePTFE membrane circumscribed by the planed root, the

- width of the alveolar bone at the apical extension of the root planing, and the membrane
3. Membrane height: distance between the apical extension of the root planing and the most coronal aspect of the ePTFE membrane
4. Junctional epithelium: distance between the apical and coronal aspect of a junctional epithelium along the planed root
5. Connective tissue repair: distance between the apical extension of the root planing and the apical extension of the junctional epithelium along the planed root
6. Cementum regeneration: distance between the apical extension of the root planing and the coronal extension of a continuous layer of new cementum or a cementum-like deposit on the planed root
7. *Bone regeneration (height)*: distance between the apical extension of the root planing and

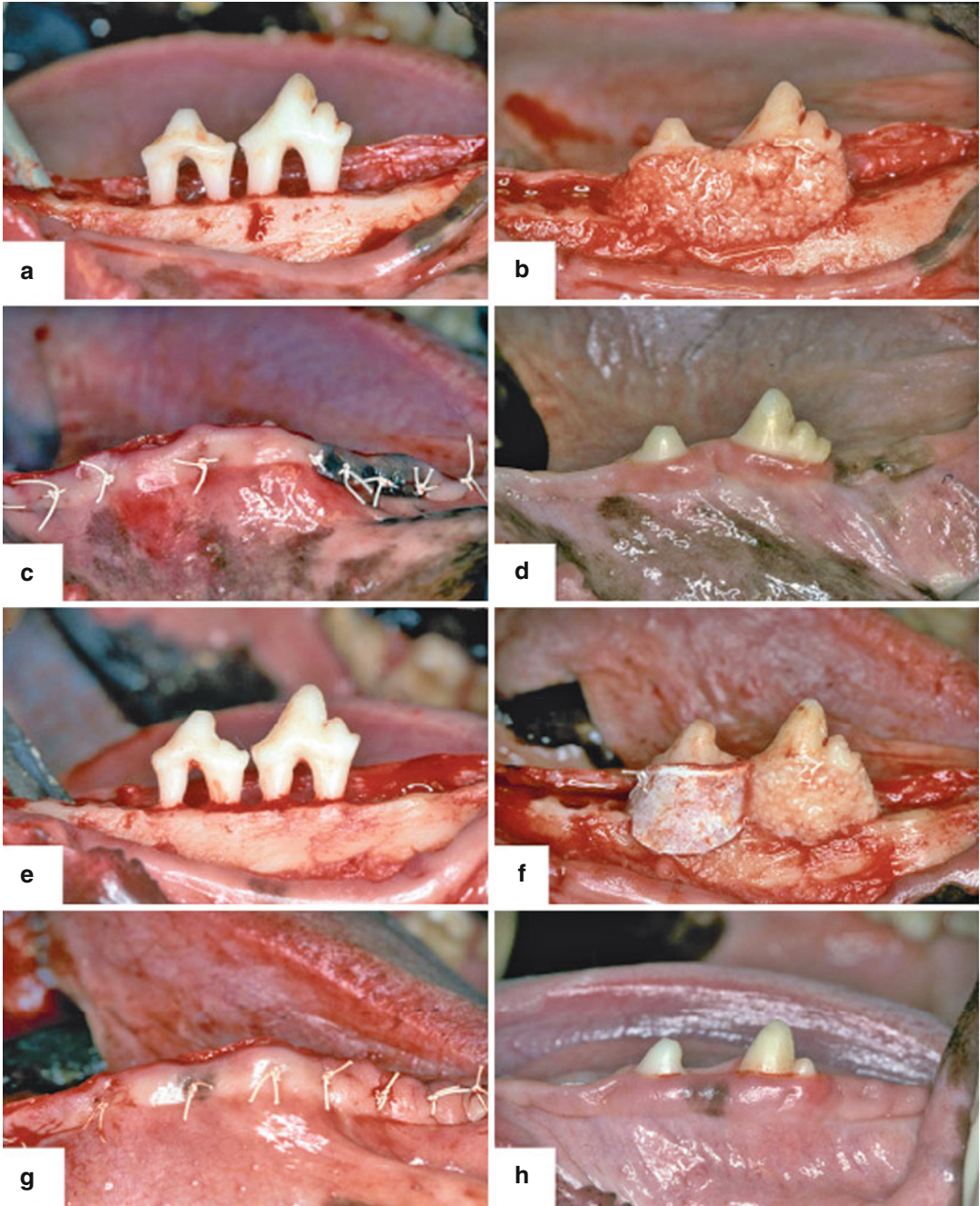


Fig. 21.6 Critical size, supra-alveolar periodontal defects were created around the third and fourth mandibular premolar teeth. The alveolar bone was removed to approximately 6 mm apical to the cemento-enamel junction of the premolar teeth (a, e). One defect site is implanted with coral implant (CI) (b), and the other site with CI/GTR (f).

Gingival flaps were advanced for transgingival wound closure (c, g). Healing at 4 weeks was generally uneventful. Exposures of the expanded polytetrafluoroethylene barrier were not observed (d, h). GTR, guided tissue regeneration (Koo et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

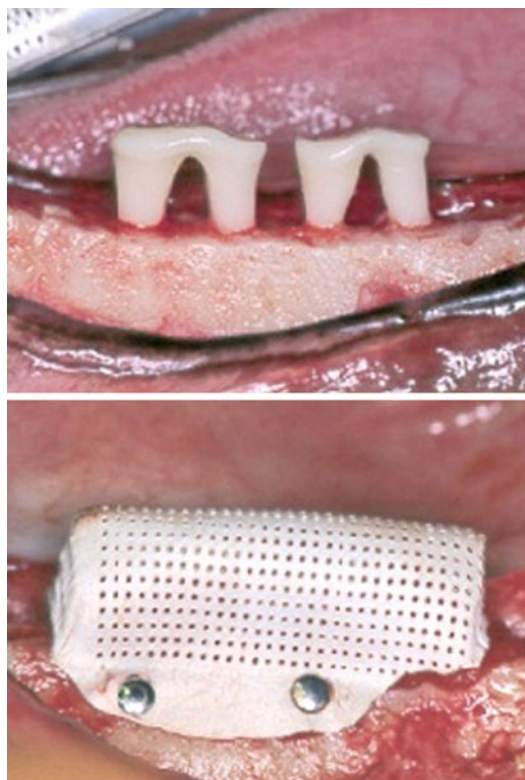


Fig. 21.7 Clinical presentation of the critical size supra-alveolar periodontal defect model, defect preparation includes circumferential surgical reduction of the alveolar bone to a level 5–6 mm apical to the cemento-enamel junction at the mandibular third and fourth premolar teeth. The periodontal attachment including the cementum is then removed by root planing leaving a distinct landmark for histometric measurements at the surgically reduced alveolar crest. The lower illustration shows installation of the space-providing, porous ePTFE device used in this study (Polimeni et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

the coronal extension of alveolar bone formation along the planed root

8. Bone regeneration (area): area represented by new alveolar bone along the planed root
9. Bone regeneration (density): ratio-mineralized bone matrix/total bone area
10. Biomaterial density: ratio residual biomaterial/total bone area
11. Root resorption: combined linear heights of distinct resorption lacunae on the planed root
12. Ankylosis: combined linear heights of ankylotic unions between new alveolar bone and the planed root

Several model systems/defects have been used to **evaluate the biologic and clinical potential of various candidate therapies for craniofacial reconstruction in general and particularly in this context alveolar augmentation**. Principally, such model systems can be divided into onlay and inlay defects (Wikesjö et al. 2006). However, it was suggested that several inlay model systems, such as tooth extraction sites (Araújo and Lindhe 2005; Araújo et al. 2005), buccal implant dehiscence defects (Hanisch et al. 2003), access defects for peri-apical surgery (Bergenholtz et al. 2006) and alveolar ridge saddle-type defects (Jovanovic et al. 2007), heal almost to completion following wound closure for primary intention healing (Figs. 21.8, 21.9, 21.10 and 21.11). Thus, the osteoconductive or osteoinductive potential of a biomaterial or other technology placed into such defects may be difficult to discern. Of course, judiciously chosen early observation intervals and/or use of fluorescent markers may provide some information whether the implanted technology supported accelerated bone formation (Wikesjö et al. 2006)

Moreover, it was demonstrated that the placement of an implant in the fresh extraction socket failed to influence the process of remodelling that occurs in the buccal and lingual hard tissue walls of the socket that followed tooth removal. Thus, after 3 months of healing, the amount of buccal bone height reduction (in comparison with lingual bone alteration) was similar at implant sites and edentulous sites. Thus, the vertical discrepancy between the buccal and lingual bone margins was in both categories of sites >2 mm: edentulous sites=2.2 mm and implant sites 2.4 mm. In other words, the resorption of the buccal bone wall of the socket was three times as great as that observed at the buccal aspect of the surgically involved control teeth (Figs. 21.12, 21.13 and 21.14) (Araújo and Lindhe 2005; Araújo et al. 2005).

Subsequently, Wikesjö et al. (2006) had modified the supra-alveolar periodontal defect model to become a relevant rigorous tool to study regeneration of alveolar bone and oral implant osseointegration, presenting the **critical-size supra-alveolar peri-implant defect model**. In

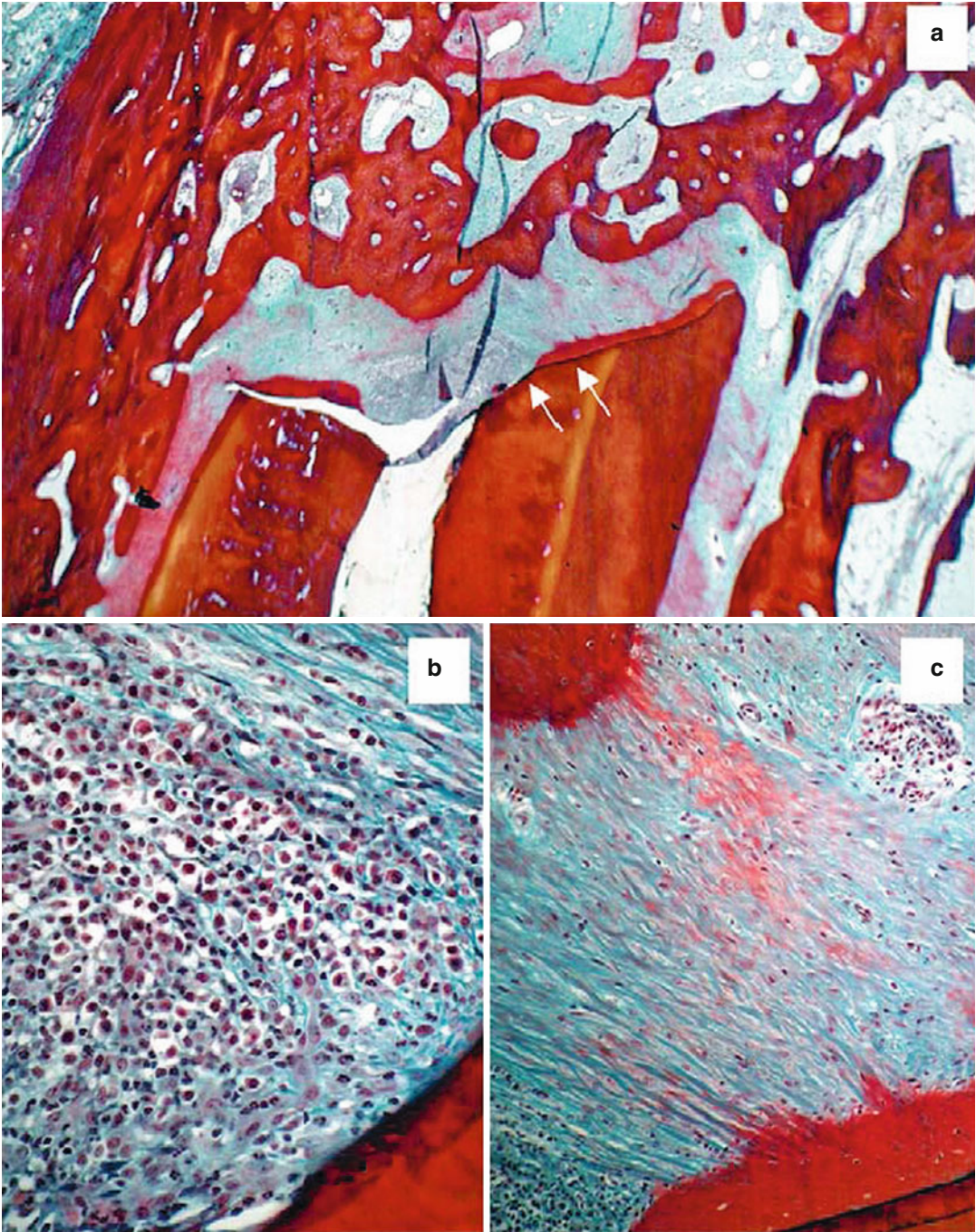


Fig. 21.8 Photomicrograph of maxillary defect site receiving sham-surgery following a 4.5-month healing interval (a). Arrows indicate areas for the higher magnifications in (b) and (c), respectively. Substantial amounts of cellular cementum have formed on the cut

dentin surface into which there is insertion of principal fibres (c). In-between the fibres near the exit of the root canal (c), there are inflammatory cell infiltrates (Bergenholtz et al. 2006. Reprinted with permission from Elsevier)

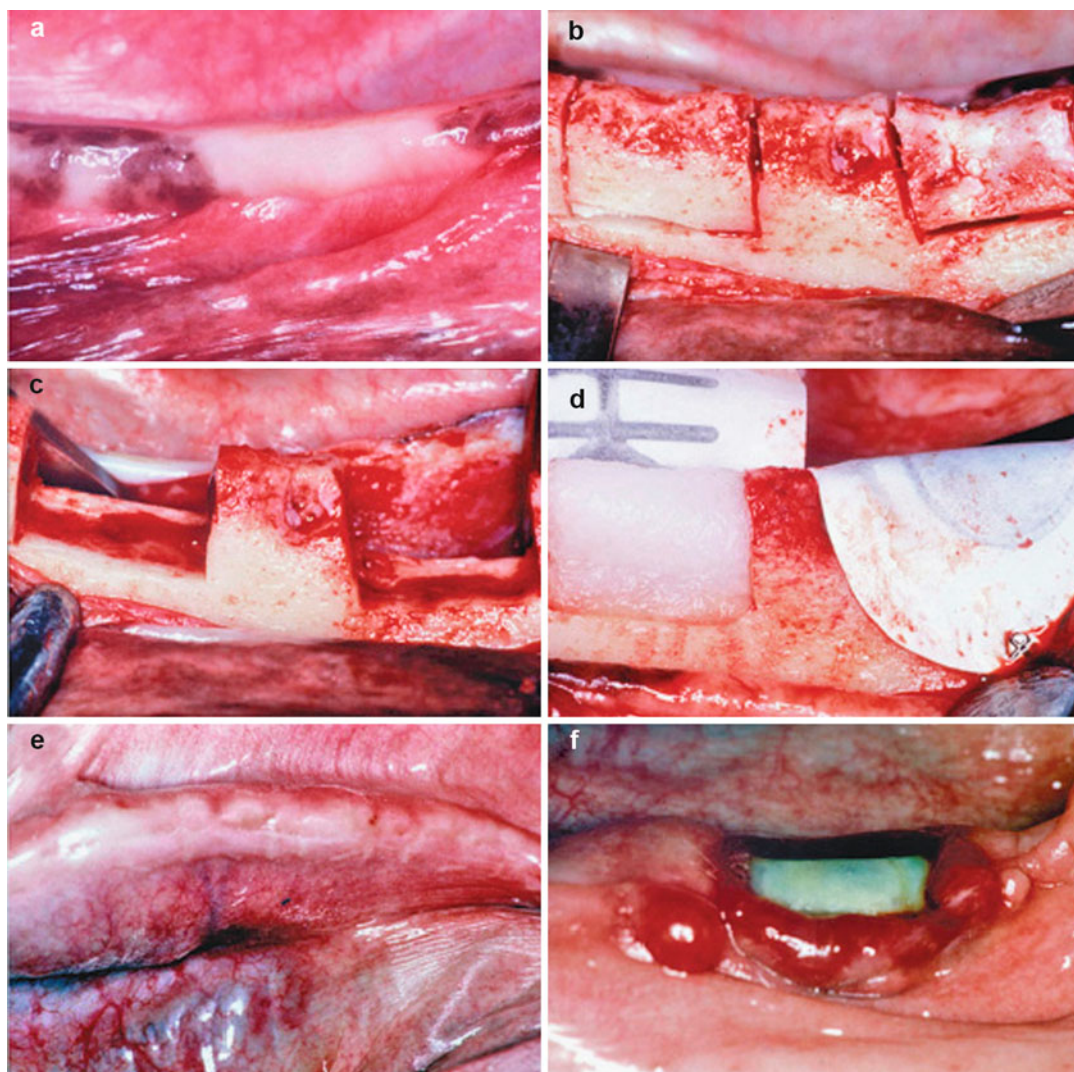


Fig. 21.9 Mandibular, alveolar ridge, saddle-type defect implanted with recombinant human bone morphogenetic protein-2 (rhBMP-2)/absorbable collagen sponge (ACS) and guided bone regeneration (GBR): pre-surgery baseline (a); surgical outline of the alveolar ridge defect (b); alveolar ridge saddle-type defect (c); application of

rhBMP-2/ACS and GBR devices (d); and clinical observations of sites implanted with rhBMP-2/ACS (e) and GBR (f). Note swelling of the site implanted with rhBMP-2/ACS and wound failure at the site receiving GBR (Jovanovic et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

brief, bilateral, critical-size, supra-alveolar, peri-implant defects were created in the mandibular premolar region (Figs. 21.15 and 21.16). Buccal and lingual mucoperiosteal flaps were reflected, and alveolar bone removed around the circumference of the premolar teeth to a level approximately 6 mm from the cemento-enamel junction using water-cooled rotating burs. The premolar teeth were then extracted, and the first molar amputated at the level of the reduced alveolar crest. Three

titanium implants were placed into osteotomies prepared into the extraction site of the distal root of the third (P1 implant) and the mesial (P2 implant) and distal (P3 implant) root of the fourth premolar in each jaw quadrant. A few implants were placed into osteotomies prepared into the reduced alveolar process when placement into extraction sites was impossible. Five millimetres of the implant was placed within the surgically reduced alveolar ridge to the level of the reference

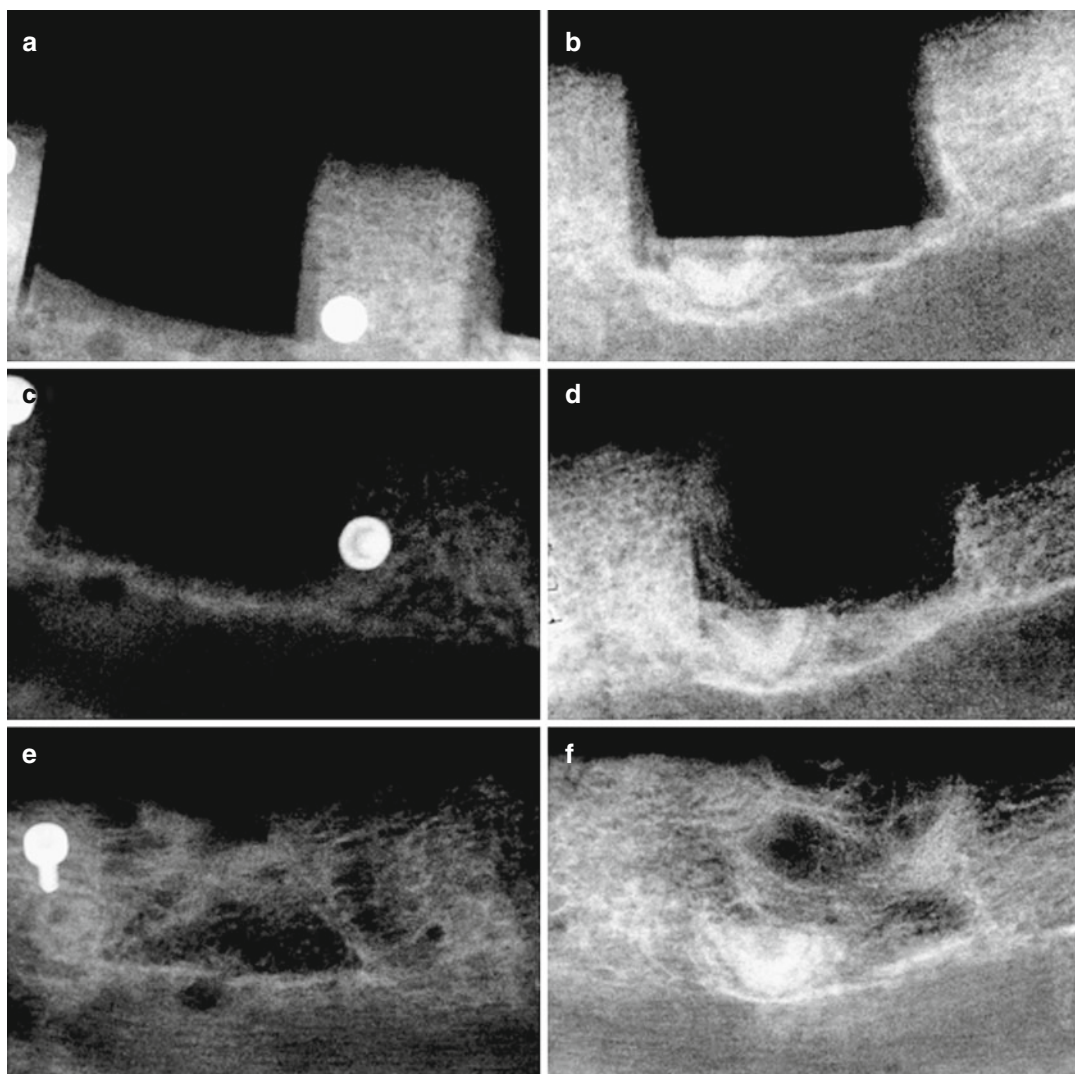


Fig. 21.10 Representative radiographs from immediately post surgery, weeks 4 and 12 post-surgery for sites receiving guided bone regeneration (GBR) (a, c, e) or recombinant human bone morphogenetic protein-2 (rhBMP-2)/absorbable collagen sponge (ACS) (b, d, f). Note bone formation from the mesial, distal, and apical

aspect of the site receiving GBR, week 4 (c). The site receiving rhBMP-2/ACS exhibits a central radiolucency indicative of seroma formation (d) that gradually resolves over the 12-week healing period (f) (Jovanovic et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

notch, creating 5 mm, supra-alveolar, peri-implant defects (Wikeshjö et al. 2006).

The most central section from each implant is used for the histometric analysis (Lee et al. 2009; Decker et al. 2010). The following measurements were recorded for the buccal and lingual surfaces of each implant (Fig. 21.17):

- *Bone regeneration (height)*: distance between the 5-mm thread and the vertical extension of newly formed bone along the implant. Bone

formation exceeding the implant platform was not included.

- *Bone regeneration (area)*: area of newly formed bone along the implant above the 5-mm thread. Bone formation exceeding the implant platform was not included.
- *Bone density (new bone)*: ratio bone/marrow spaces in newly formed bone.
- *Osseointegration (new bone)*: per cent bone–implant contact (BIC) measured between the

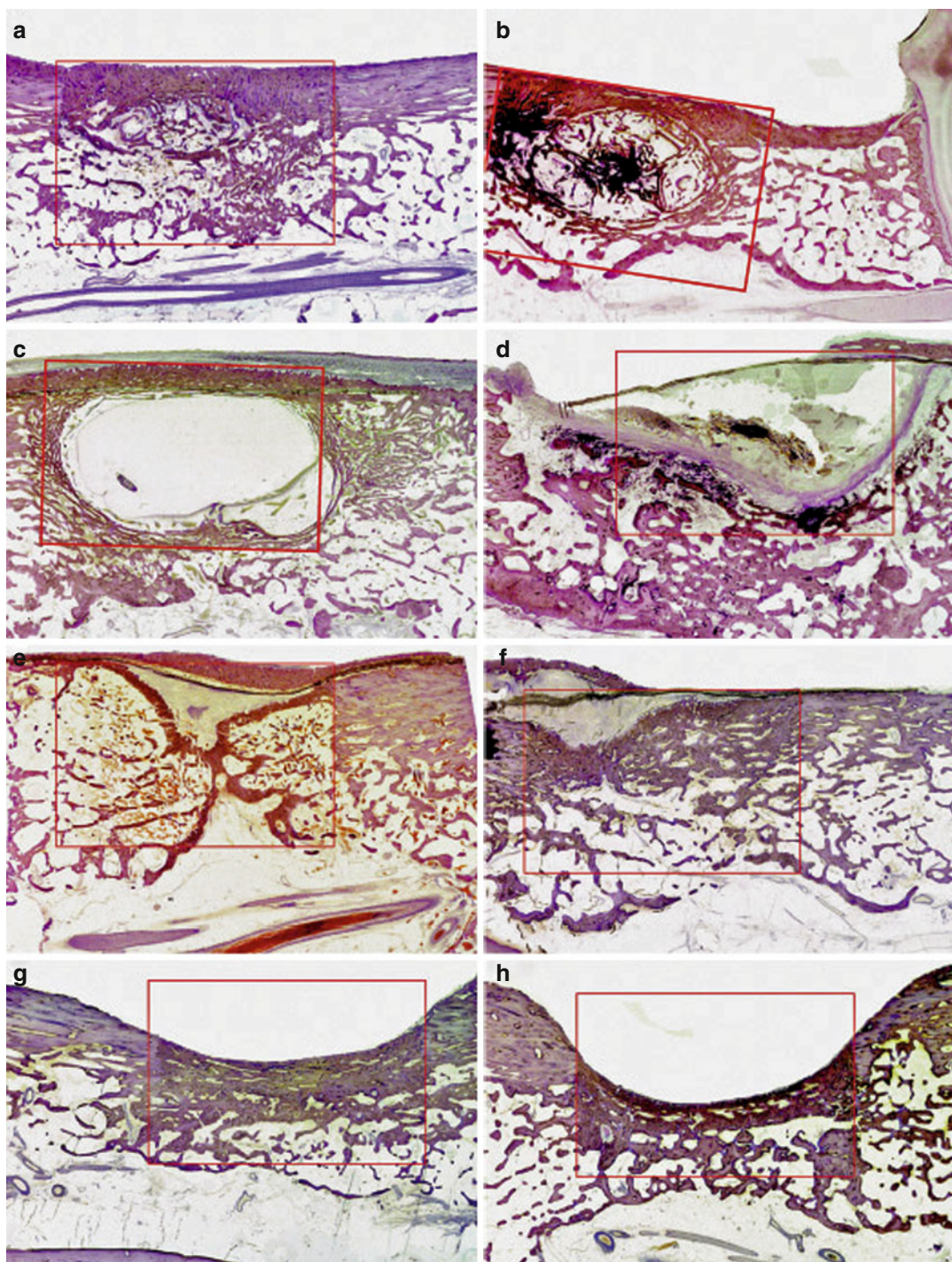


Fig. 21.11 Representative photomicrographs of defect sites receiving recombinant human bone morphogenetic protein-2 (rhBMP-2)/absorbable collagen sponge (ACS) ((a) cortex formation and complete trabecular bone fill; (b) cortex formation and resolving seroma filled with trabecular bone); rhBMP-2/guided bone regeneration (GBR) ((c) cortex formation and large seroma; (d) wound failure/

membrane exposure, note cortex formation over part of the GBR barrier); GBR ((e) cortex formation; (f) limited, late (?) wound failure/membrane exposure, note cortex formation over part of the GBR barrier); and surgery control with (g) or without (h) ACS. Red frames approximate the original defect sites (Jovanovic et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

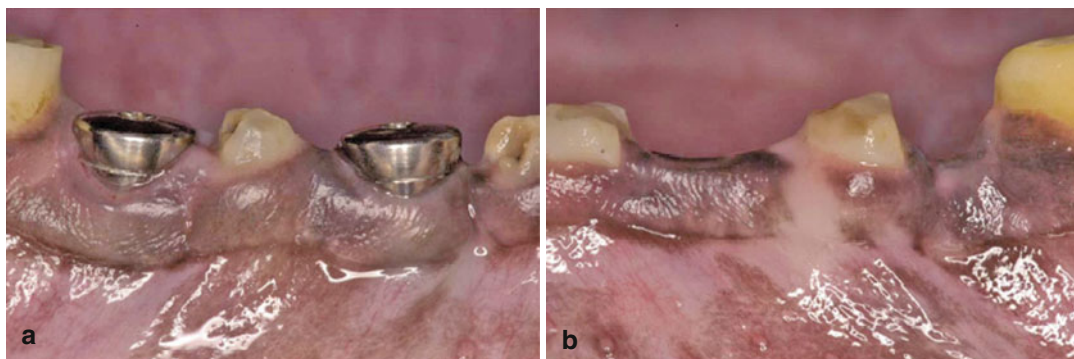


Fig. 21.12 (a) Clinical photograph illustrating one experimental site (two implants and two 'involved' roots) after 3 months of healing. Note that the peri-implant mucosa as well as the gingiva shows no overt signs of inflammation. The margin of the mucosa resides at the

smooth portion of the implant. (b) Clinical photograph of two edentulous sites and adjacent 'involved' tooth sites after 3 months of healing (Araújo et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

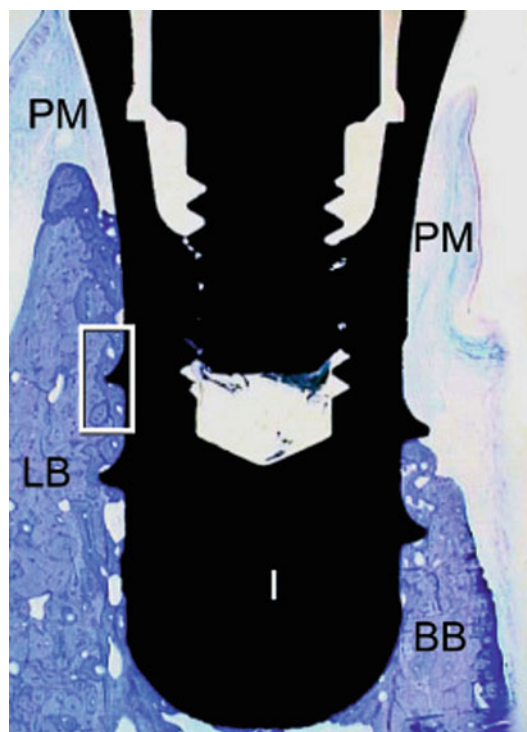


Fig. 21.13 Buccal-lingual section representing one implant site after 3 months of healing. Note the location of the bone crest at the buccal and lingual aspects of the implant. BB buccal bone wall, I implant, LB lingual bone wall, PM peri-implant mucosa. Toluidine blue staining; original magnification 16× (Araújo et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

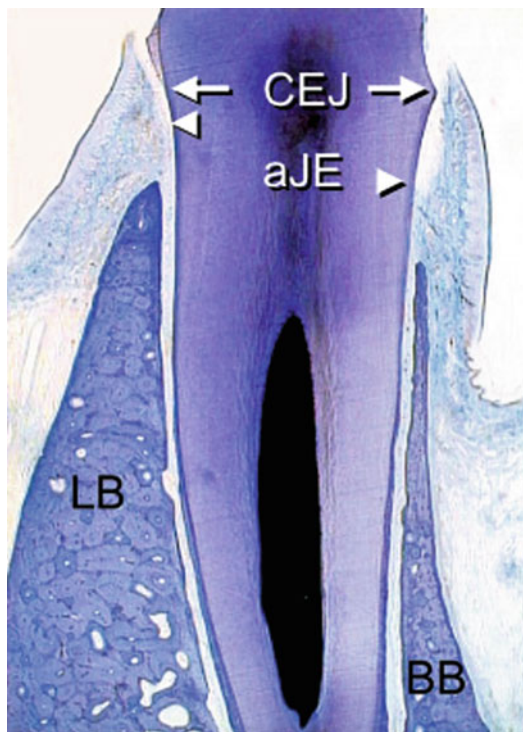


Fig. 21.14 Buccal-lingual section representing an involved tooth site. Note that the lingual bone crest is closer to the CEJ (arrows) at the lingual than at the buccal aspect of the tooth. The apical level (aJE) of the junctional epithelium (arrowheads). BB buccal bone wall, LB lingual bone wall, CEJ cemento-enamel junction. Toluidine blue staining; original magnification 16× (Araújo et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

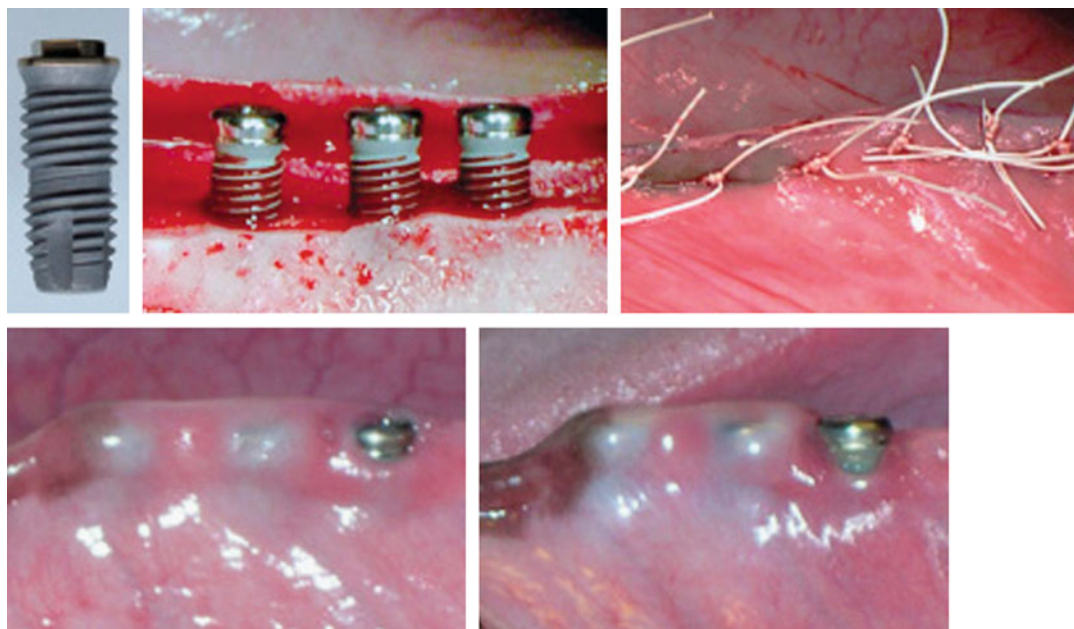


Fig. 21.15 Custom-produced 10 mm implant with a notch at the 5 mm level. The clinical pictures show the surgically reduced alveolar ridge with implants placed into the extraction site of the distal root of the mandibular third premolar (*left*; P1 implant), or the mesial (P2 implant) or distal (P3 implant) root of the fourth premo-

lar, following wound closure, and at weeks 4 and 8 post surgery. The contours of the implants can be discerned through the oral mucosa, the most coronal aspect of the P3 implant becoming exposed at week 4 (Wikesjö et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

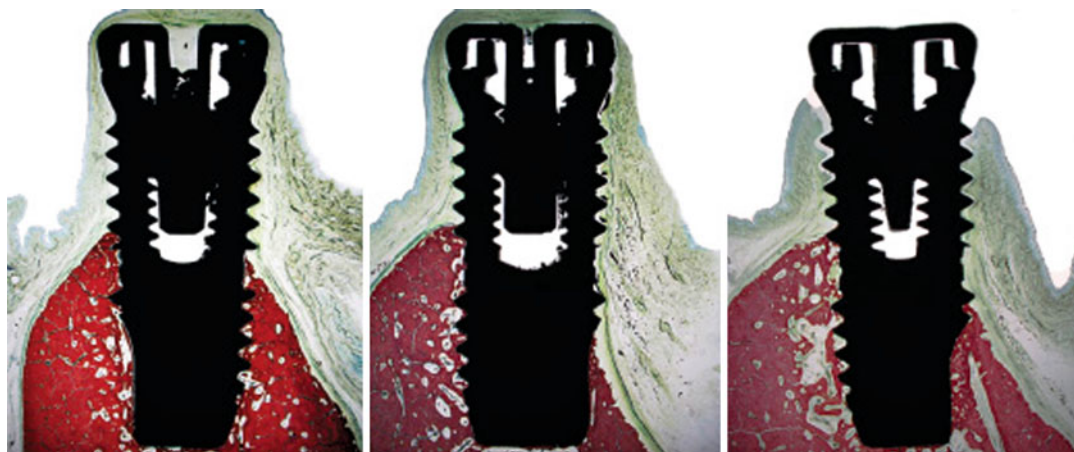


Fig. 21.16 Photomicrographs showing buccal-lingual sections of the implants in **Fig. 21.1** (P1–P3 implant; *left/right*) at week 8 post-surgery. Note that the lingual (*left*) aspect of the implants exhibits limited new bone for-

mation whereas the buccal aspect shows bone resorption, in particular for the P2 and P3 implants; Stevenel's blue/van Gieson's picro fuchsin (Wikesjö et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

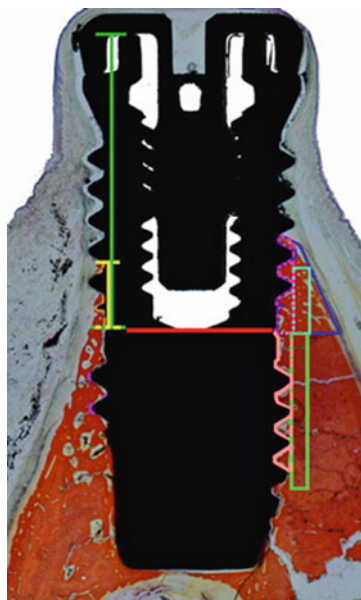


Fig. 21.17 Schematic representation of the histometric analysis. Reference thread (red line). Implant left side: defect height (green line), bone height (yellow line), new bone–implant contact (BIC) (dashed orange line) and resident BIC (dashed pink line). Implant right side: bone area (blue enclosure), new bone density outside the implant threads (light blue box), new bone density within the implant threads (pink enclosure), resident bone density within the implant threads (pale pink enclosure), and resident bone density outside the implant threads (green box) (Lee et al. 2009). Reprinted with permission from John Wiley & Sons, Inc.)

5-mm thread and the vertical extension of newly formed bone along the implant.

- *Bone density outside the implant threads/BD_{OT}* (resident bone): ratio bone/marrow spaces in a 400 × 4,000-μm area (width × height) immediately outside the implant threads in resident bone.
- *Bone density within the implant threads/BD_{WT}* (resident bone): ratio bone/marrow spaces within the implant threads in resident bone.
- *Osseointegration (resident bone)*: per cent BIC within resident bone measured from the 5-mm thread to the apex of the implant.

The critical-size, supra-alveolar, peri-implant defect model, a genuine onlay defect model, has limited innate osteogenic potential under optimal conditions for healing, and this model appears a rigorous and preferred tool in the critical evaluation of candidate technologies including bone biomate-

rials, devices for GBR, and implantable or injectable technologies using growth and differentiation factors for alveolar regeneration and osseointegration of endosseous oral implants (Wikesjö et al. 2006). High examiner reproducibility and temporal stability can be achieved for histometric data acquisition using the critical-size supra-alveolar peri-implant defect model. Examiner reproducibility should be routinely assessed, reported and accounted for to assure the quality of evidence generated by preclinical studies (Lee et al. 2009).

The critical-size supra-alveolar periodontal defect and subsequently the critical-size supra-alveolar peri-implant defect model, developed for the assessment of periodontal and alveolar bone regenerative technologies, are well-characterized canine models that provide a discriminating evaluation of candidate therapies that have been successfully screened in laboratory bench evaluations and small animal model systems. These defect models have proven to be a ‘litmus test’ for candidate therapies for periodontal wound healing/regeneration and alveolar bone augmentation/osseointegration, respectively (Lee et al. 2009).

21.4 Nonhuman Primates

All nonhuman primates have a deciduous and a permanent dentition; the dental formula of the permanent dentition in baboons (*Papio anubis*) and cynomolgus monkeys (*Macaca fascicularis*) is reported to be the same as in humans (I 2/2, C 1/1, Pm 2/2, and M 3/3), whereas that of the cotton-eared marmosets (*Callithrix jacchus*), cotton-top marmosets (*Saguinus oedipus*), squirrel monkeys, and bushbabies are different (Schou et al. 1993).

Macroscopically, tooth morphology, occlusion, and gingiva in baboons, howler monkeys (*Alouatta caraya*), cynomolgus and squirrel monkeys, and cotton-top and cotton-eared marmosets are quite similar to humans with the following exceptions: larger canines, presence of diastemata between anterior teeth primarily at the canine, and edge-to-edge relationship of the incisors. This indicates that molar and premolar regions are to be preferred for periodontal investigations in which the diastema is not desirable. However,

the size of the dentition deviates substantially from that of humans (Schou et al. 1993).

Clinically, healthy periodontal of cotton-top marmosets, baboons, bushbabies and stump-tailed squirrel and howler monkeys has been observed to be histologically quite similar to that of humans. It is characterized by a non-keratinized junctional and oral sulcular epithelium without or with minimal rete ridges, spares but varying accumulation of inflammatory cells in the connective tissue and highly organized bundles of collagen fibres in the gingival connective tissue and periodontal ligament (Fig. 21.18) (Schou et al. 1993).

21.4.1 Naturally Occurring and Experimental Periodontitis in Nonhuman Primates

There are four types of periodontal defects in nonhuman primates, which have been used. These include defects produced by naturally occurring periodontitis and three types of experimentally produced defects: the acute defect model, the chronic defect model and the acute/chronic defect model (Caton et al. 1994).

The acute defect model is produced by the following procedures: mucoperiosteal flaps are elevated, and the bone, periodontal ligament and cementum are removed to create the desired defect shape. Flaps are either replaced and suture or the experimental variable is inserted prior to flap closure (Fig. 21.19) (Caton et al. 1994). The acute fenestration-type defect model has often been used for investigating the early healing events following different types of therapeutic procedures. In this defect model, the defect space is completely surrounded by residual periodontal tissue, allowing progenitor cells to repopulate the defect from all sides. Furthermore, the location of the defect prevents flap dehiscence with subsequent inflammation and infection, thus ensuring an optimal wound stability. Consequently, this defect model is not comparable with the regular chronic periodontal defect but offers some advantages in investigating possible differences in the quality of the newly

formed tissues, by minimizing the negative influence of flap dehiscence or infection (Sculean et al. 2000a).

The problem with the acute defect model is that spontaneous regeneration occurs. Advantages of the acute wound healing model include cost-effectiveness and reduced experimental time (Caton et al. 1994).

Chronic periodontal defects are created in monkeys by placing orthodontic elastics around the circumference of teeth or slightly apical to the gingival margin (Fig. 21.20) (Kornman et al. 1981a, 1982; Siegrist and Kornman 1982; Kiel et al. 1983; Manti et al. 1984; Brex et al. 1985; Brex and Patters 1985; Ebersole et al. 2000, 2010; Oates et al. 2002; Kostopoulos and Karring 2004; Branch-Mays et al. 2008; Cappelli et al. 2009). The elastic gradually migrates apically as plaque-induced inflammation destroys the supporting periodontal ligament and bone. Advanced pocketing and bone loss is achieved in 3–6 months. The elastics migrate faster on single-rooted teeth and those with more tapered roots. The advantages of the chronic defect model are that severe defects are created which behave as human periodontal defects. Spontaneous regeneration is not observed, and the response to flap procedures is the formation of a long junctional epithelium.

A disadvantage is that the defects take up to 6 months to produce, and, therefore, this model is expensive to use in experiments due to the cost of prolonged animal care. Furthermore, the deep defects produced by orthodontic elastics are limited to the interproximal region. The defects produced on the facial and lingual surfaces are much shallower. This may make this model inappropriate for certain types of experiments (Caton et al. 1994).

Kornman et al. (1981a) had characterized the microbial and tissue changes following ligature placement in *M. fascicularis* and divided the disease process into four stages: Stage I—a chronic gingivitis condition without discernible bone loss prior to ligature in which Gram-positive rods account for the majority of organisms; stage II—development within 1–3 weeks of a flora consisting of Gram-negative rods, especially *P.*



Fig. 21.18 Characteristic clinical, radiographic and histological features of implants and teeth with healthy peri-implant mucosa/gingiva (**a**, **b**) or peri-implantitis/

periodontitis (**c**, **d**) in monkeys (Schou et al. 2002. Reprinted with permission from John Wiley & Sons, Inc.)



Fig. 21.19 The standardized fenestration defect created at the vestibular aspect of tooth 23 following surgical removal of the buccal bone and of the root cementum (Sculean et al. 2000a. Reprinted with permission from Springer)



Fig. 21.20 Experimental tooth with an artificial composite crown during the period of ligature-induced periodontitis (Kostopoulos and Karring 2004. Reprinted with permission from John Wiley & Sons, Inc.)

gingivalis and *Capnocytophaga*; stage III—occurs at 4–7 weeks after ligature placement. The microbial population is comprised of Gram-negative rods with *P. gingivalis* accounting for over one-third of the isolated organisms. This stage coincided with initiation of bone resorption and pocket formation; and stage IV—at 8–17 weeks—the flora shifts and appears to exhibit characteristic microorganisms intermediate to stages II and III. *C. sputigena* predominated stage IV lesions, with anaerobic streptococci comprising the largest Gram-positive coccal population, and anaerobic bacteria in general remaining at a level greater than the facultative organisms.

It was recently demonstrated that nonhuman primates with high levels of long-standing gingival inflammation when compared to those nonhuman primates with low inflammation show (1) different inflammatory mediator profiles in gingival crevicular fluid (particularly for immunoglobulin A (IgA) and IgG levels), (2) a different quantitative and qualitative subgingival microbiota and (3) a similar progression of periodontitis. Thus, while variations in host inflammatory responses to local factors exist in the nonhuman primates, an extensive subgingival challenge (such as ligation) may negate these individual differences (Ebersole et al. 2000).

The acute/chronic defect model is developed by surgically creating the periodontal defect in the desired location and shape by removing bone, periodontal ligament, and cementum. Prior to flap closure, the defect is encouraged to enter a chronically inflamed state by placement of a foreign object, such as ligatures. The advantages of the acute/chronic defect model are many. Defects are produced rapidly, and, therefore, the model is more cost-effective due to reduced animal care expenses. Bilateral defects can be made identical relative to bone and connective tissue loss. Finally, the defects do not heal with significant amounts of spontaneous regeneration (Caton et al. 1994).

The use of nonhuman primates was later modified to include **inoculation with human pathogens** (Oz et al. 2010). The successful implantation of a rifampin-resistant strain of the putative periodontal pathogen *Bacteroides gingivalis* (*P. gingivalis*) into the periodontal microbiota of monkeys (*Macaca fascicularis*) resulted in an increase in the systemic levels of antibody to the microorganism and rapid and significant bone loss (Holt et al. 1988).

The histological manifestations of **spontaneous gingivitis and periodontitis** in baboons, bushbabies, chimpanzees and howler, cynomolgus and squirrel monkeys and cotton-top and cotton-eared marmosets are quite similar to those of humans. Characteristic features are progressive rete ridge formation, ulceration, apical migration of the pocket epithelium with wide intercellular spaces and emigrating PMNs, increased infiltration of inflammatory cells in

connective tissue, vascular proliferation, destruction of collagen fibres and resorption of alveolar bone. Also, the histological features of experimentally produced lesions in rhesus, squirrel and cynomolgus monkeys are similar to periodontal lesions in humans. Lymphocytes and plasma cells dominate the inflammatory infiltrate of cynomolgus monkeys, while the inflammatory infiltrate of squirrel monkeys was different, with a predominance of PMNs and macrophages and very few plasma cells. Even through most of these studies in only 14-day studies and/or studies of the connective tissue apically to the pocket epithelium, it was concluded that squirrel monkeys appear poorly related to the human condition (Schou et al. 1993).

21.4.2 Summaries of Studies Using Nonhuman Primates in Periodontal Research

In the last 10 years, 25 articles have been published using monkeys for research relating to periodontal healing, filling with biomaterials, guided tissue regeneration, enamel matrix derivatives or implant surgery (Struillou et al. 2010). These surgical approaches are, for the most part, carried out on *Macaca fascicularis*. All the teeth can be used, which makes it possible to obtain an important number of test sites, with a limited number of animals (Table 21.2). For example, in the study by Sculean et al. (2000b), 24 lesions (intraosseous and fenestration defects) were created at the level of the incisors and premolars in only three monkeys. The intraosseous defects ($n=18$) were located in the mesial part of each tooth and had a depth of between 6 and 8 mm. In order to prevent any spontaneous healing and to promote plaque accumulation, metal strips were placed in the defects and were attached to the teeth using a composite material. Standardized critical-size defects of the fenestration type ($n=6$) were made in the same monkeys. The defects were treated by guided tissue regeneration (GTR) or by induced guided regeneration by enamel matrix proteins (EMD). Control defects were treated with coronally repositioned

flaps. This model is quite similar to those used in clinical human studies and, for this reason, appears to be a pertinent animal model (Struillou et al. 2010).

For histometric assessments, the following linear distances can be measured as follows (Fig. 21.21): (1) free gingival margin to the apical termination of the dentogingival epithelium (PD), that is, pocket depth; (2) level of crown amputation to the apical termination of the dentogingival epithelium (AL), that is, loss of connective tissue attachment; (3) level of crown amputation to the crest of the bone defect (CBL), that is, the coronal bone level; (4) level of crown amputation to the bottom of the bone defect (ABL), that is, the apical bone level; (5) apical border of the notch in the root surface to the coronal level of newly formed cementum (NC), that is, the amount of newly formed attachment (only on test roots); and (6) apical border of the notch to the coronal level of newly formed bone (BR), that is, the amount of newly formed bone (only on test roots) (Kostopoulos and Karring 2004).

21.4.3 Advantages and Disadvantages for Using Nonhuman Primates in Periodontal Research

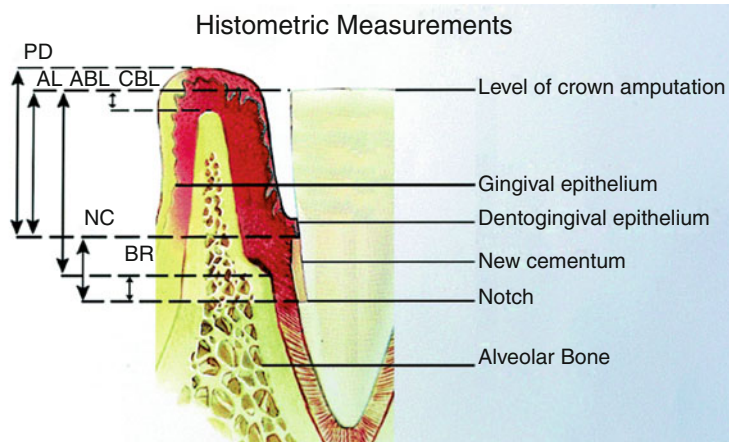
The nonhuman primate has the main advantage of possessing anatomical relationships closest to the human and potentially adequate numbers of available class II furcations (Giannobile et al. 1994).

Disadvantages of the monkey include the low prevalence of natural disease, difficulties of handling, and the low frequency of class III furcations (Giannobile et al. 1994). Although various species of nonhuman primates are adequate for studying different aspects of periodontal diseases, monkeys are expensive to purchase and maintain and are ferocious. Wild captured monkeys can be disease carriers. *Mycobacterium tuberculosis*, *Herpes simplex type B*, *Shigella* species and Simian B virus are infectious to man. Monkeys are also prone to systemic infectious diseases and pose difficulties in controlling post-surgical

Table 21.2 Selection of recent studies using non-human primates in periodontal research (Persson 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

Author, year	Aim of the study	Design	Primate species	Animals	Defect type	Tooth	Evaluation
Sculean et al. (1997)	Periodontal healing	Split-mouth	Macaca fascicularis	1	Intrabony (bone defects, with a depth of approximately 6–8 mm and a width in the bucco-lingual direction of 3–5 mm)	Incisor, premolar and molar	Histomorphometry
Hammarström et al. (1997)	EMD	Split-mouth	Macaca fascicularis	29	Buccal dehiscence	Canine, premolar and molar	Histomorphometry
Sculean et al. (1998)	EMD; GTR; EMD+GTR	Split-mouth	Macaca fascicularis	3	One-wall intrabony, furcation III defect, fenestration	Incisors, canine, premolar and molar	Histomorphometry
Blumenthal et al. (2003)	Collagen membrane (GTR), human demineralized freeze-dried bone (DFDB) grafting (BG), combined therapy (GTR + BG) and a DFDB–glycoprotein sponge matrix (MAT)	Split-mouth	Baboons (Papio anubis)	9	Intrabony Three walls Two walls	Incisors, canine, premolar and molar	Clinical parameters Histomorphometry Periodontal regeneration
Kostopoulos and Karring (2004)	GTR (e-PTFE membrane)	Split-mouth	Macaca fascicularis	4	Ligature-induced periodontitis	Incisor, premolar and molar	Histomorphometry
Sculean et al. (2000a)	GTR, EMD, or control	Split-mouth	Macaca fascicularis	3	Fenestration	Canine	Histomorphometry
Sculean et al. (2000b)	GTR, EMD, GTR + EMD	Split-mouth	Macaca fascicularis	3	Intrabony defects + metal strips placed into the defects to enhance plaque accumulation	Incisor, premolar and molar	Histomorphometry
Donos et al. (2003)	GTR, EMD, GTR + EMD	Split-mouth	Macaca fascicularis	3	Furcation III defect	Molar	Histomorphometry
Graziani et al. (2005)	GTR	Parallel	Macaca fascicularis	10	Buccal dehiscence-type defects	Premolar and molar	Histomorphometry
Sculean et al. (2001)	GTR, EMD, GTR + EMD	Split-mouth	Macaca fascicularis	3	Intrabony defects + metal strips placed into the defects to enhance plaque accumulation, fenestration	Incisor, canine, premolar	Immunohistochemistry

Fig. 21.21 Schematic illustration of histometric assessments (Kostopoulos and Karring 2004. Reprinted with permission from John Wiley & Sons, Inc.)



infections and trauma (Weinberg and Bral 1999). Small nonhuman primates like squirrel monkeys and marmosets, although less expensive and easier to cage than macaque species, develop a different periodontal disease histopathology, with few to no lymphocytes and plasma cells. These primates may also be difficult to manipulate because of their small size. Therefore, the use of squirrel monkeys and marmosets may not be appropriate in many studies of periodontal disease pathogenesis (Schou et al. 1993).

21.4.4 Nonhuman Primate Study Models in Vaccine Trials Against Periodontitis

Nonhuman primates, including *M. fascicularis*, *M. nemestrina*, marmosets, baboons and chimpanzees, have been considered for periodontal vaccine trials. Results from active immunization studies using an experimental periodontitis model in *M. fascicularis* are presented (Table 21.3). Detailed assessments of immunization impact with regard to immunological, microbiological and clinical outcomes are summarized. The results from these studies demonstrate that the relationship between antibody titres and killing abilities and protection against challenge with *P. gingivalis* infection in the nonhuman primate model is complex (Persson 2005).

21.5 Dogs

Many experimental studies on gingival and periodontal diseases have been conducted in dogs. The beagle is one of the most commonly used due to its size and its extremely cooperative temperament. Globally, all periodontal tissues and the size of the teeth are quite similar to those observed in humans (Fig. 21.22). However, some major differences exist between dogs and humans as the lack of lateral movements, no occlusal contacts for all the premolars, and presence of open contacts between teeth. The frequent lack of gingival sulci and crevicular fluid, a different composition of periodontal plaque and calculus are other important differences between dogs and humans (Struillou et al. 2010). The subgingival microflora in beagle dogs with relatively healthy gingiva is different from humans with a high % of Gram-negative bacteria (Kornman et al. 1981b).

Permanent teeth of the dog consist of three incisors, one canine, four premolars and three molars in the mandible and two molars in the maxilla (Weinberg and Bral 1999). The study demonstrated that it is possible in dogs to establish and maintain a normal gingiva simply by eliminating calculus and then subjecting the animals to daily repeated and carefully performed tooth cleanings (Lindhe et al. 1975). Juvenile dog gingiva seems to differ in the following respects from adult dog gingiva: a thicker keratinized layer of the oral epithelium, a junctional epithelium structurally

Table 21.3 Examples of non-human primate study models and results from vaccine trials against periodontitis (Persson 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Antigen	Study type	Results	Conclusion
Moritz et al. (1998)	Purified cysteine protease (porphy-pain-2) from <i>P. gingivalis</i> versus placebo immuniza-tion. For experi-mental protocol, see Ebersole et al. (1991)	Case-control study of experimental periodontitis in <i>M. fascicularis</i> Clinical measures, and radiographic analysis (CADIA) ELISA assays	1. Elevated serum IgG titres to whole-cell <i>P. gingivalis</i> (36-fold) and Porphy-pain-2 (194-fold) 2. 25% more Gram-negative bacteria at control sites than in immunized animals 3. Few clinical changes as an effect of immunization 4. No statistically significant changes as defined by CADIA between sham- or test-immunized animals	Immunization with porphy-pain-2 induces an immune response 1 Immunization may impact microbial colonization Clinical effect remains unclear
Persson et al. (1994a)	Vaccine composed of formalin-killed whole-cell <i>P. gingivalis</i> (5083, primate strain) and (Syntex SAF) adjuvant	Case-control study of 10+10 active/sham immunized <i>M. fascicularis</i> with pre-existing low IgG titres but presence of <i>P. gingivalis</i> in pockets over 44 weeks with infectious challenge at week 36. Immunization at baseline weeks 3, 6, and 16 when ligatures were placed. Routine clinical measures, standardized radiographs (CADIA), ELISA assays, and DNA probe analysis	1. Serum IgG titre to <i>P. gingivalis</i> elevation through immunization with 50% titre levels remaining at endpoint 2. Trend towards less <i>P. gingivalis</i> in immunized animals 3. Significantly more bone loss in control animals at ligated sites and exaggerated by microbial challenge 4. Variety in antibody response	Immunization with formalin-killed whole-cell <i>P. gingivalis</i> with an adjuvant inhibits progres-sion of experi-mental periodontitis
Persson et al. (1994b)	DNA probe analysis of the microflora in relation to clinical features	Cross-section study of the oral microflora in <i>M. fascicularis</i> Routine clinical periodontal measures	1. <i>A. actinomycetemcomi-tans</i> , <i>P. gingivalis</i> , <i>P. intermedia</i> , <i>C. rectus</i> , <i>T. forsythia</i> <i>F. nucleatum</i> present in majority of sites tested in non-experimental conditions 2. Antibody titre levels to <i>P. gingivalis</i> inversely correlated with <i>P. gingivalis</i> levels	<i>M. fascicularis</i> appears to be a suitable model in which studies of vaccine efficacy for control of periodontitis can be tested
Ebersole et al. (1990)	<i>A. actinomycetem-comitans</i> leucotoxin produced in a vector DNA model was used to immunize <i>M. fascicularis</i>	Development and testing of a leucotoxin vaccine derived from <i>A. actinomycetem-comitans</i> and tested in <i>M. fascicularis</i> to assess serum responses Case-control animal study	1. Most <i>M. fascicularis</i> carried antibody titres to <i>A. actinomycetemcomitans</i> 2. Primary immunization elicited Ig1 and Ig3 responses 3. Secondary response elicited IgG2 responses and increased antibody avidity	A host response to immunization with <i>A. actinomy-cetemcomitans</i> leucotoxin with potential efficacy possible
Nisengard et al. (1989)	<i>B. macacae</i> (non-human primate equivalent to <i>P. gingivalis</i> in humans)	Preliminary case-control study in <i>M. fascicularis</i> Ligature-induced periodon-titis 12-week immunization period Clinical measures, GCF radiographs Serum titres to <i>B. macacae</i>	1. Immunization-induced elevated antibody IgG titres to <i>B. macacae</i> 2. Ligatures-induced bone loss in all animals 3. Levels <i>B. macacae</i> were 2 times higher in non-immu-nized animals after 6 months	Heightened immune responses and partial prevention of microbial colonization

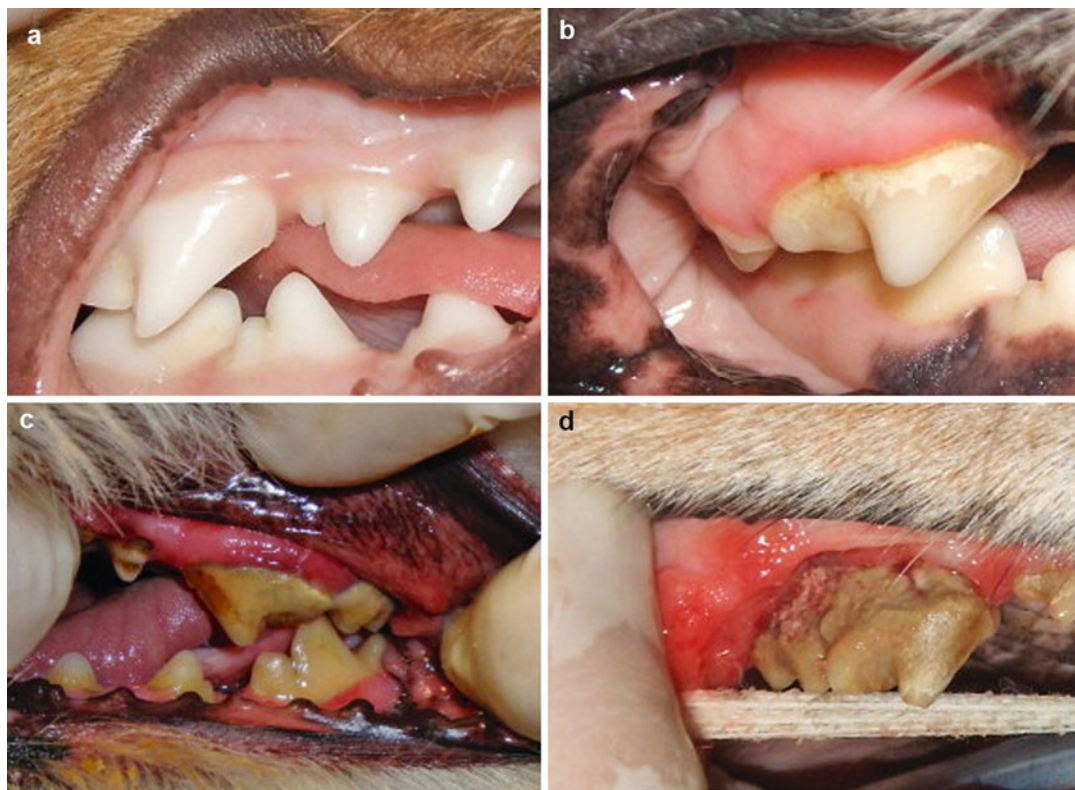


Fig. 21.22 Periodontium at different stages of health/disease. **(a)** Healthy gingiva; photograph of the maxillary and mandibular right arcades of a dog showing normal gingival tissues. **(b)** Mild gingivitis; photograph of the maxillary right fourth premolar of a dog showing a slight gingival margin erythema and mild calculus on the teeth. **(c)** Moderate gingivitis; photograph of the maxillary left

fourth premolar of a dog showing a significant erythema and oedema to the gingival, as well as increasing calculus accumulation on the tooth. **(d)** Advanced periodontitis; photograph of the maxillary right fourth premolar of a dog showing an obvious gingival recession, severe gingival inflammation and calculus accumulation (Albuquerque et al. 2012. Reprinted with permission from Elsevier)

resembling the oral epithelium, a cuticular structure at the surface of the junctional epithelium, a limited mononuclear inflammatory cell response during experimental gingivitis and a delayed establishment of an infiltrated connective tissue portion during experimental gingivitis. These differences might explain the lower propensity of juvenile dog gingiva to react with clinically notable inflammation during a period without oral hygiene. In addition, during experimental gingivitis, subgingival plaque formed along the tooth surface to a lesser extent in juvenile compared to adult dogs. Therefore, there seem to be both bacterial factors and host factors which may be of importance for the establishment of gingival inflammation in juvenile dogs (Matsson and Attstrom 1979).

21.5.1 Naturally Occurring and Experimental Periodontitis in Dogs

It has been demonstrated that *induced gingivitis* in young beagle dogs can be induced by withdrawing oral hygiene for 7 days in animals in which healthy gingiva had previously been maintained (Soames and Davies 1980). Differences exist between dog and man in the location of the inflammatory infiltrate in early gingivitis. In the dog, the initial infiltrate located in the marginal part of the gingiva, proceeds along the junctional epithelium leaving the connective tissue in a relatively normal state. In healthy dogs kept plaque-free for prolonged periods of time, a gingival

sulcus is not evident, but a pocket develops with the onset of gingivitis (Lindhe and Rylander 1975; Weinberg and Bral 1999). It was also observed that plaque (and later calculus) formed more rapidly on posterior premolars and molars than on anterior premolars and incisors. This indicates that in dogs fed a soft diet, gross plaque formation must be limited by factors such as the anatomy of the oral cavity and the friction produced during mastication and by movements of the tongue and cheek. These differences in the rate of plaque and calculus formation in different tooth regions also resulted in a different rate of deterioration of the periodontal tissues in various parts of the beagle dog dentition (Lindhe et al. 1975).

Originally, it was reported that periodontal disease began slowly in young dogs and increased with age, progressing about 5× faster in dogs than humans. Later, it was documented that the range and severity of gingivitis and periodontitis varied in both young and older dogs and that gingivitis in younger dogs did not necessarily progress into periodontitis. Gingival recession is an outstanding feature in dogs with periodontitis (Weinberg and Bral 1999 and references therein). Disease progression in the canine with **natural periodontal disease** coincides with distinct supragingival and subgingival microbial populations. The supragingival plaque found coronal to the gingival tissues contains primarily Gram-positive rods and cocci, especially *A. viscosus*. The subgingival plaque as found in periodontitis lesions consists predominantly of anaerobic Gram-negative organisms with significant levels of *P. gingivalis*, *F. nucleatum* and *Capnocytophaga* (Giannobile et al. 1994). Spontaneous class II furcation defects occurring with periodontal disease in old dogs were used in guided tissue regeneration research (Caffesse et al. 1990; Dyer et al. 1993; Lekovic et al. 1998; Elharar et al. 1998).

Kornman et al. (1981b) indicated that **ligature placement** (silk bindings around the teeth for a period of 4–6 months) in beagle dogs is associated with changes in the subgingival microflora which were characterized by an increase in the propositional levels of *B. asaccharolyticus*, including catalase-positive *B. asaccharolyticus*-like organisms, forming 34.7% of the cultivable flora. A shift in the flora from Gram-negative facultative rods to

Gram-negative anaerobic rods was also observed. Lindhe and Ericsson (1978) had demonstrated that the application of cotton floss ligatures around the neck of the lower premolars of dogs fed a diet which favours plaque formation and establishes conditions which promote development of gingivitis and rapid destruction of supporting apparatus. The histometric assessments also showed that the ligatures also play a role not only for the initiation but also for the maintenance of rapidly advancing periodontitis in the dog. Thus, measurements made in biopsies sampled 180 days after ligature placement revealed that on average, 32.5% of the fibre attachment to the root surface was lost. Assessments made in tissue sections sampled 360 days after ligature placement showed that on average, 46.1% of the supporting apparatus has been lost. It has been suggested that the ligature operates as a collector for dental plaque and provides favourable living conditions for plaque bacteria. It was also reasonable to assume that the ligature favours the accumulation of large amounts of subgingival plaque.

Various types of defect models in dogs with different topologies have been established, including supra-alveolar, dehiscence, fenestration, furcation, 1-wall, 2-wall, 3-wall and interproximal defect models.

Experimentally produced furcation defects in the dog were described by Araujo and Lindhe (1998) and Araujo et al. (1999). ‘Through and through’ furcation defects, about 4 mm high and 3 mm wide, were produced on third lower premolars after the second and fourth mandibular premolars were extracted (Fig. 21.23). Reconstructive surgery was subsequently performed in group A using a GTR technique. In this context, it should be realized that in contrast to an intrabony defect or a traditional extraction socket, the furcation defect in the current experiment was located in a ‘suprabony’ position and had only one bone wall. It was noted that a fibrous-rich provisional connective tissue formed and completely filled the open furcation space already at the 4-week interval. The supra-alveolar furcation defect became between the 1- and 2-month interval filled with a primary spongiosa, made up mainly of woven bone. The amount of mineralized bone and bone marrow that was present in the furcation defect

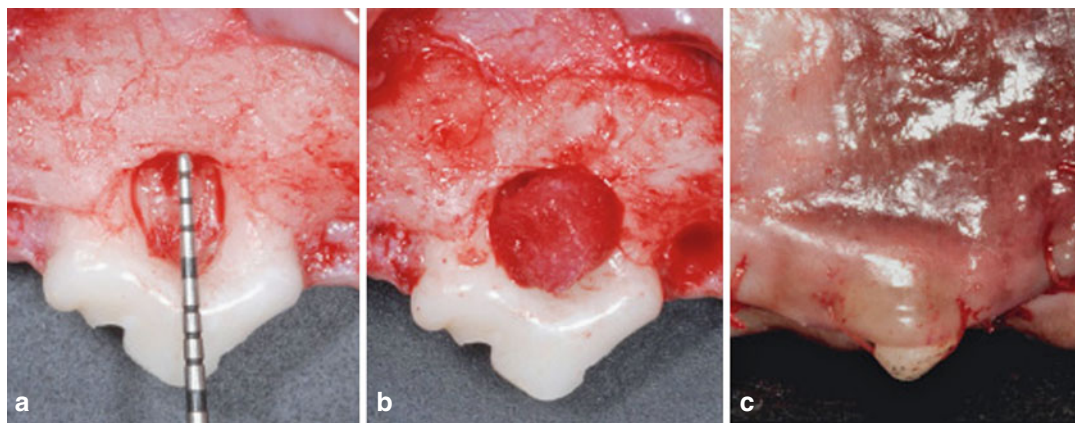


Fig. 21.23 Clinical illustration of treatment procedure: (a) aspect of the Class II furcation defect after the defect creation. (b) The collagen sponge scaffold, with (test) or without (control) the cell suspension was applied filling

the defect. (c) Flaps coronally positioned and sutured after membrane positioning (Suaïd et al. 2011). Reprinted with permission from John Wiley & Sons, Inc.)

varied with time. Thus, the proportion of mineralized bone decreased while the proportion of bone marrow increased consistently from the second to the fifth month of healing. This indicates that an intense modelling activity occurred during this 3-month period. This modelling resulted in the formation of (1) one large marrow space in the centre of the bone tissue and, in addition, (2) a smaller bone marrow space in the most coronal portion of the furcation. At the end of the study (6 months), the bone marrow occupied a much larger space than in the bone tissue of pristine furcations (Araujo et al. 1999). The model was recently used also for assessing periodontal healing by autologous periodontal ligament cells (Suaïd et al. 2011; Dogan et al. 2002).

The **critical-size, supra-alveolar, periodontal defect model** was first presented in 1990 (Wikesjö and Nilvéus 1990; Wikesjö et al. 1994) and has been used extensively to evaluate biologic and environmental factors influencing periodontal wound healing/regeneration as well as candidate regenerative protocols, including root surface conditioning, devices for guided tissue regeneration (GTR), bone biomaterials, and extracellular matrix, growth and differentiation factors (for a review, see Polimeni et al. 2006; Wikesjö and Selvig 1999). As Struillou et al. (2010) presented, two methodological approaches were described as follows: treatment of surgical defects on the third and fourth mandibular premolars (**Fig. 21.24**)

(Koo et al. 2004b; Wikesjö et al. 2003a, b) or a test based only on the third premolar after avulsion of the second and fourth premolars (**Fig. 21.23**) (Araujo et al. 1998, 1999; Araujo and Lindhe 1998; Roriz et al. 2006). The osteogenic potential in this defect model following sham surgery amounts to less than 20% of the defect height and a 4- or 8-week healing interval, suggesting that bone formation may be completed within 4 weeks. Similarly, cementum regeneration amounts to less than 10% of the defect height following a 4- or 8-week healing interval. This challenging model system has been proven valuable for qualitative and quantitative evaluation of candidate protocols for periodontal regeneration (Koo et al. 2004b). The supra-alveolar periodontal defect model particularly lends itself to histometric evaluations of regeneration of alveolar bone, cementum, a functionally oriented periodontal ligament, formation of a junctional epithelium, position of any implanted device and position/amount of any residual biomaterial used in conjunction with the regenerative procedure, as well as qualitative and quantitative evaluations of root resorption and ankylosis (Koo et al. 2004b). However, the supra-alveolar defect model is difficult to prepare because the entire coronal alveolar bone around the tooth must be removed. In addition, there is a large difference between intrabony and supra-alveolar defects in their degrees of difficulty in spontaneous healing (Jung et al. 2011).



Fig. 21.24 The critical-size, supra-alveolar periodontal defect model. The defect height from the reduced alveolar bone to the cemento-enamel junction approximates 6 mm (Koo et al. 2004b. Reprinted with permission from John Wiley & Sons, Inc.)

Surgically created intraosseous defects with one wall (critical-size, ‘box-type’, 4×4×4-mm periodontal defect at the distal aspect of the second and mesial aspect of the fourth mandibular premolar teeth) (Kim et al. 2002, 2004, 2005; Song et al. 2005; Shirakata et al. 2007; Sakata et al. 2006; Stavropoulos and Wikesjö 2010; Min et al. 2011), two walls (Ozmeric et al. 2000; Kim et al. 2004; Felipe et al. 2009), and three walls (4×4×4 mm), surgically created at the mesial and distal aspect of the maxillary and mandibular first and third premolars, respectively (Kim et al. 1998, 2004; Choi et al. 2002), were used in preclinical regenerative therapy.

A comparison of one-, two-, and three-wall intrabony defects has shown that two-wall intrabony defects result in a larger coefficient of variance in histometric analysis than one- and three-wall intrabony defects. This is due to different defect morphology and tissue resources. The location/selection of the histological section in the two-wall intrabony defect may contribute considerably to this large variability. Thus, two-wall intrabony defects are not the preferred model for evaluation of periodontal regeneration therapies in experimentally induced lesions. Only one- and three-wall defects appear to be reproducible models (Tables 21.4, 21.5 and 21.6). However, because of the superior measurement outcomes, the one-wall rather than the three-wall intrabony defect should be considered when quantitative evaluations are required (Kim et al. 2004, 2011).

To evaluate the influence of defect dimensions on periodontal wound healing/regeneration in intrabony defects, contralateral one-wall intrabony [6×6 mm (wide/deep) vs. 4×4 mm (narrow/shallow)] (Figs. 21.25 and 21.26) periodontal defects were surgically created at the edentulated mesial aspect of the mandibular first molars in dogs. Both wide/deep and narrow/shallow intrabony defects showed a substantial potential for periodontal regeneration in this preclinical model (Stavropoulos and Wikesjö 2010).

Limited periodontal regeneration following gingival flap surgery alone following an 8-week healing interval has been reported for the intrabony model system, alveolar bone, and cementum regeneration not exceeding 30–35% of the defect height. Although the **8-week histological observations** may not fully explain early healing events, an 8-week healing interval has been considered useful to study periodontal repair/regeneration in dog models (Fig. 21.27) (Kim et al. 2005). Moreover, Choi et al. (2002) observed no additional bone and cementum formation between an 8- and 24-week healing interval for sham-surgery controls in a dog intrabony defect model suggesting longer observation intervals may not be necessary to capture the osteogenic potential and cementogenesis (Kim et al. 2005).

The maxillary teeth of the dog were considered unsuitable in a periodontal defect model because of the dissimilarity of the dog and human palates, especially compared with the similarity in the mandibular teeth. Hence, most periodontal defect models were prepared at the mandibular premolar area. However, because dogs have four premolars and two molars in the maxilla compared with three molars in the mandible, the space between the premolars is suitable for preparing a proximal defect (Jung et al. 2011). An **interproximal periodontal defect (IPD)**, which can be categorized as a no-wall defect that faces root to root, could be ranked between these two defect models: intrabony and supra-alveolar. The configuration of IPD defect in clinical situation is complicated and is often encountered in daily practice. The validity of the IPD in dogs was recently demonstrated. Surgery was performed in the interproximal area between the maxillary second and third premolars in two beagle dogs.

Table 21.4 One-wall intrabony defects (Kim et al. 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Author, year	Animal	Defect size (mm)	Healing interval (weeks)	Materials	Defect site (mm)	Bone (mm)	Cementum (mm)	Connective tissue adhesion (mm)	Junctional epithelium (mm)
Park et al. (1998)	Dog	4 × 4	8	Bioactive glass Control	4.0 ± 0.3 4.2 ± 0.5	2.4 ± 0.6 1.8 ± 0.7	2.6 ± 0.4 2.1 ± 0.7	0.2 ± 0.3 0.4 ± 0.2	1.1 ± 0.4 ^a 1.7 ± 0.5
Kim et al. (2002)	Dog	4 × 4 × 4	8	Safflower seed extract/collagen Phosphate-buffered saline/collagen Control	4.0 ± 0.2 4.1 ± 0.2 4.0 ± 0.1	2.9 ± 0.7 ^a 2.1 ± 0.6 1.2 ± 0.7	3.8 ± 0.6 ^a 3.8 ± 0.2 ^a 1.5 ± 1.2	0.3 ± 0.4 ^a 1.1 ± 1.6 ^a 1.4 ± 0.3	-0.6 ± 0.7 -0.1 ± 0.4 1.1 ± 1.7
Park et al. (2003)	Dog	4 × 4	8	Chitosan/collagen sponge Buffer control/collagen sponge Control	4.1 ± 0.1 4.0 ± 0.1 4.2 ± 0.1	2.4 ± 0.4 ^a 1.5 ± 0.4 1.0 ± 0.8	3.5 ± 0.8 ^a 1.6 ± 0.4 1.4 ± 0.5	0.4 ± 0.4 1.1 ± 0.9 0.7 ± 0.6	0.3 ± 0.6 ^a 1.5 ± 1.3 2.3 ± 1.2
Yeo et al. (2005)	Dog	4 × 4 × 4	8	Chitosan non-woven membrane Resorbable membrane Control	4.9 ± 0.2 4.9 ± 1.1 4.5 ± 0.6	1.8 ± 0.2 ^a 1.5 ± 0.7 1.2 ± 0.3	2.3 ± 0.2 ^a 1.7 ± 0.6 1.4 ± 0.2	1.1 ± 0.2 1.2 ± 0.4 0.9 ± 0.3	1.8 ± 0.7 2.0 ± 0.6 2.3 ± 0.3
Baik et al. (2004)	Dog	4 × 4	8	Calcium phosphate glass CaP + resorbable membrane Control	4.6 ± 0.7 4.9 ± 0.6 4.8 ± 0.5	43.5 ± 13.3% 36.5 ± 15.1% 27.2 ± 7.5%	49.2 ± 12.1% ^a 39.6 ± 12.1% 32.9 ± 10.5%	27.0 ± 4.2% 27.9 ± 9.7% 36.4 ± 9.0%	24.1 ± 9.1% 38.7 ± 12.2% 30.9 ± 9.9%
Song et al. (2005)	Dog	4 × 4 × 4	8	Safflower seed extract/bioabsorbable membrane Bioabsorbable membrane Control	4.8 ± 0.2 4.7 ± 0.3 4.3 ± 0.4	2.6 ± 0.7 ^a 2.4 ± 0.3 ^a 1.7 ± 0.3	3.7 ± 0.8 ^a 3.2 ± 0.4 ^a 2.5 ± 0.4	1.0 ± 0.2 0.9 ± 0.1 ^a 1.2 ± 0.2	0.2 ± 0.7 0.5 ± 0.5 0.6 ± 0.8

Min et al. (2005)	Dog	4 × 4	12	Chitosan non-woven membrane	4.4 ± 0.4	28.3 ± 14.8%	30.7 ± 14.6%	10.7 ± 5.1% ^a	58.6 ± 13.1%
				Chitosan non-woven membrane + CaP block bone	4.9 ± 0.6	52.3 ± 14.3% ^a	55.3 ± 14.3% ^a	11.4 ± 5.1% ^a	33.2 ± 13.2%
				Control	4.8 ± 0.9	30.4 ± 4.1%	26.9 ± 13.9%	28.6 ± 13.9%	44.5 ± 14.5%
Kim et al. (2007)	Dog	4 × 4	8	Poly(lactic acid)/polyglycolic acid membrane	5.1 ± 0.4	2.4 ± 0.5 ^a	3.2 ± 0.4 ^a	0.7 ± 0.2	1.2 ± 0.3
				Poly(lactic acid)/polyglycolic acid/tetracycline membrane	5.0 ± 0.5	2.9 ± 0.7 ^a	3.7 ± 0.5 ^a	0.6 ± 0.1	0.7 ± 0.2 ^a
				Control	4.3 ± 1.1	1.5 ± 0.7	2.0 ± 0.7	0.9 ± 0.6	1.5 ± 0.3
Kim et al. (2009)	Dog	4 × 5	8	rhGDF-5 (0 µg)/ACS	4.5 ± 0.7	1.4 ± 0.4	2.5 ± 0.7	0.1 ± 0.1	1.9 ± 0.8
				rhGDF-5 (1 µg)/ACS	4.9 ± 0.7	2.1 ± 0.8 ^a	3.5 ± 0.9 ^a	0.1 ± 0.1	1.3 ± 0.8
				rhGDF-5 (20 µg)/ACS	5.0 ± 0.5	2.2 ± 0.8	3.0 ± 1.2	0.4 ± 0.9	1.6 ± 1.2
				rhGDF-5 (100 µg)/ACS	4.6 ± 0.7	2.2 ± 0.4 ^a	3.5 ± 0.8 ^a	0.1 ± 0.1	1.1 ± 0.7

ACS absorbable collagen sponge, *rhGDF-5* recombinant human growth and differentiation factor-5

^aSignificantly different from controls ($P < 0.05$)

Table 21.5 Two-wall intrabony defects (Kim et al. 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Author, year	Animal	Defect size (mm)	Healing interval	Materials	Defect site (mm)	Bone (mm)	Cementum (mm)	Connective tissue adhesion (mm)	Junctional epithelium (mm)
Ozmeric et al. (2000)	Dog	5 × 4 × 5–6	30 days	Collagen membrane	4.8 ± 0.2	1.8 ± 0.1		1.3 ± 0.1	1.7 ± 0.2
				Control	5.3 ± 0.1	2.2 ± 0.2		1.4 ± 0.1	1.2 ± 0.1
			60 days	Collagen membrane	4.8 ± 0.2	2.2 ± 0.2		2.2 ± 0.2 ^a	2.1 ± 0.2
				Control	5.6 ± 0.2	2.7 ± 0.2		1.5 ± 0.1	1.7 ± 0.1
			90 days	Collagen membrane	5.7 ± 0.3	2.5 ± 0.3		2.2 ± 0.2	2.2 ± 0.2
Blumenthal et al. (2003)	NHP	3 × 5	6 weeks	Control	5.1 ± 0.3	1.9 ± 0.1		2.8 ± 0.1	2.1 ± 0.1
				Collagen membrane	4.6 ± 1.0	1.3 ± 0.7		1.5 ± 0.4	1.0 ± 0.4
				Human DBM	4.5 ± 0.8	1.7 ± 1.0		0.8 ± 0.4	1.9 ± 1.0
				Human DBM + collagen membrane	4.4 ± 0.8	1.9 ± 0.4		1.7 ± 0.6	1.3 ± 0.9
				Human DBM + glycoprotein matrix	4.5 ± 0.2	1.8 ± 0.8		1.0 ± 1.0	1.4 ± 1.0
Villaca et al. (2005)	NHP	2 × 5	50 days	Bioactive glass		3.0 ± 0.4			1.3 ± 0.1
				Control		1.5 ± 0.9			2.8 ± 0.5
			90 days	Bioactive glass		2.5 ± 0.4			2.1 ± 0.4
				Control		1.0 ± 1.4			3.3 ± 0.7
Yamamoto et al. (2007)	Dog	5 × 5 × 5	8 weeks	EMD + bovine bone		52.2 ± 21.7%	89.8 ± 11.7%		21.7 ± 1.4%
				Bovine bone		31.6 ± 9.6%	45.0 ± 18.1%		20.6 ± 1.9%
				Control		18.2 ± 2.51%	16.2 ± 1.29%		55.3 ± 0.9%

DMB demineralized bone, EMD enamel matrix derivatives, NHP non-human primates

^aSignificantly different from control (*P* < 0.05)

Table 21.6 Three-wall intrabony defects (Kim et al. 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

Author, year	Animal	Defect size (mm)	Healing interval (weeks)	Materials	Defect site (mm)	Bone (mm)	Cementum (mm)	Connective tissue adhesion (mm)	Junctional epithelium (mm)
Kim et al. (1998)	Dog	4 × 4 × 4	8	DBM + Ca ₂ SO ₄	4.2 ± 0.5	2.7 ± 0.4 ^a	3.0 ± 0.3 ^a	0.4 ± 0.3 ^a	0.6 ± 0.3
				DBM	4.3 ± 0.7	2.7 ± 0.3 ^a	3.1 ± 0.4 ^a	0.4 ± 0.3 ^a	0.5 ± 0.2
				Ca ₂ SO ₄	4.0 ± 0.2	1.8 ± 0.5 ^a	2.5 ± 0.4 ^a	0.5 ± 0.2 ^a	1.0 ± 0.5
				Sham-surgery control	4.1 ± 0.2	0.7 ± 0.1	1.6 ± 0.3	1.6 ± 0.5	0.9 ± 0.5
Choi et al. (2002)	Dog	4 × 4 × 4	8	rhBMP-2/ACS	5.5 ± 0.3	3.6 ± 0.6 ^a	3.0 ± 1.4	1.6 ± 1.5	0.8 ± 0.7
				Buffer control	5.0 ± 0.4	2.6 ± 0.6 ^a	2.5 ± 1.5	1.5 ± 1.6	1.0 ± 0.5
				Control	5.0 ± 0.3	2.3 ± 0.4	2.5 ± 1.2	1.5 ± 1.2	1.0 ± 0.5
		24		rhBMP-2/ACS	5.0 ± 0.4	3.4 ± 0.6 ^a	2.4 ± 1.2	1.9 ± 1.2	0.7 ± 0.2
				Buffer control	4.9 ± 0.7	2.6 ± 0.6 ^a	2.7 ± 1.7 ^a	1.4 ± 1.5	0.9 ± 0.5
				Control	4.8 ± 0.4	2.1 ± 0.6	1.6 ± 1.3	2.1 ± 1.8	1.1 ± 0.6
Shirakata et al. (2002)	Dog	4 × 4 × 8	12	Injectable CaP cement	6.7 ± 0.9	5.8 ± 1.5	6.0 ± 1.3	6.7 ± 1.1	
Blumenthal et al. (2003)	NHP	3 × 5	6	Control	7.5 ± 0.8	5.1 ± 1.2	5.9 ± 0.9	6.3 ± 1.0	
Onodera et al. (2005)	Dog	6–8	2	Collagen membrane	4.3 ± 0.1	2.0 ± 0.8		2.2 ± 1.0	0.9 ± 0.3
				Human DBM	4.4 ± 0.1	1.9 ± 0.6		1.0 ± 0.4	1.8 ± 0.9
				Emdogain® + GTR		24.9 ± 0.1	19.7 ± 0.1		
				GTR		10.0 ± 0.01	9.3 ± 0.1		
Hayashi et al. (2006)	Dog	4 × 5 × 5	12	Emdogain® + GTR		84.0 ± 0.1	77.6 ± 0.2		
				GTR		57.2 ± 0.1	55.2 ± 0.1		
				Emdogain® + GTR		99.0 ± 0.1	98.3 ± 0.04		
				GTR		95.9 ± 0.1	90.1 ± 0.12		
Hayashi et al. (2006)	Dog	4 × 5 × 5	12	Injectable CaP cement	7.03 ± 0.31	4.90 ± 0.56 ^a	6.00 ± 0.36 ^a	6.46 ± 0.34 ^a	
				Control	6.71 ± 0.31	3.63 ± 0.68	4.57 ± 0.62	5.14 ± 0.58	

ACS absorbable collagen sponge, GTR guided tissue regeneration, rhBMP-2 recombinant human bone morphogenetic protein-2

^aSignificantly different from control (*P* < 0.05)

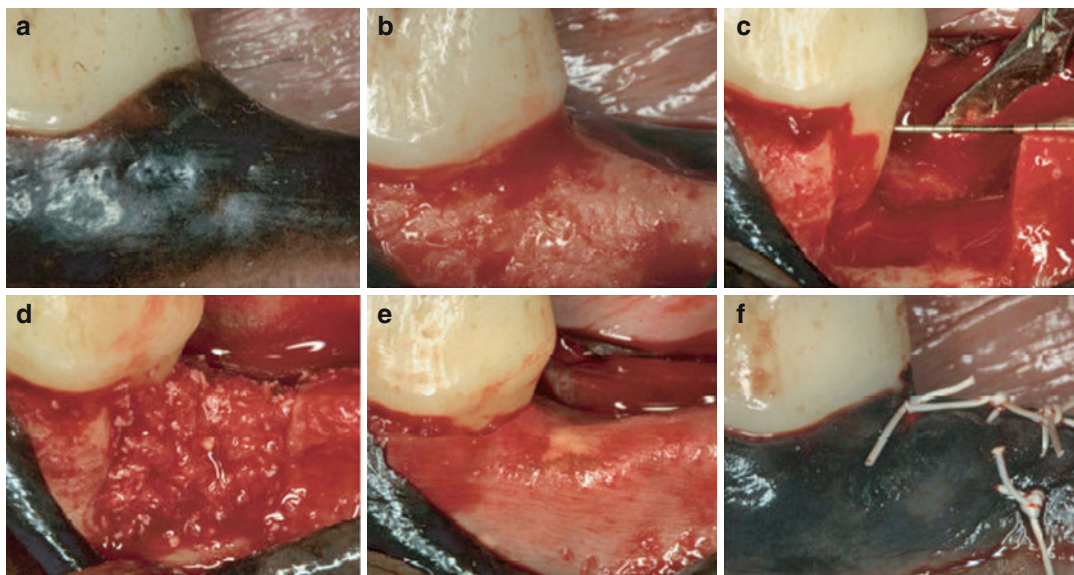


Fig. 21.25 Pre-surgery view of the experimental area (a), after flap elevation (b), and creation of a box-type, wide/deep (6×6 mm) intrabony periodontal defect at the mesial aspect of the mandibular first molar (c). The defect was filled with a piece of a bovine bone/collagen matrix

(d), covered with a porcine Type I collagen membrane (e) and the flaps were repositioned and sutured for primary intention healing (f) (Stavropoulos and Wikesjö 2010. Reprinted with permission from John Wiley & Sons, Inc.)

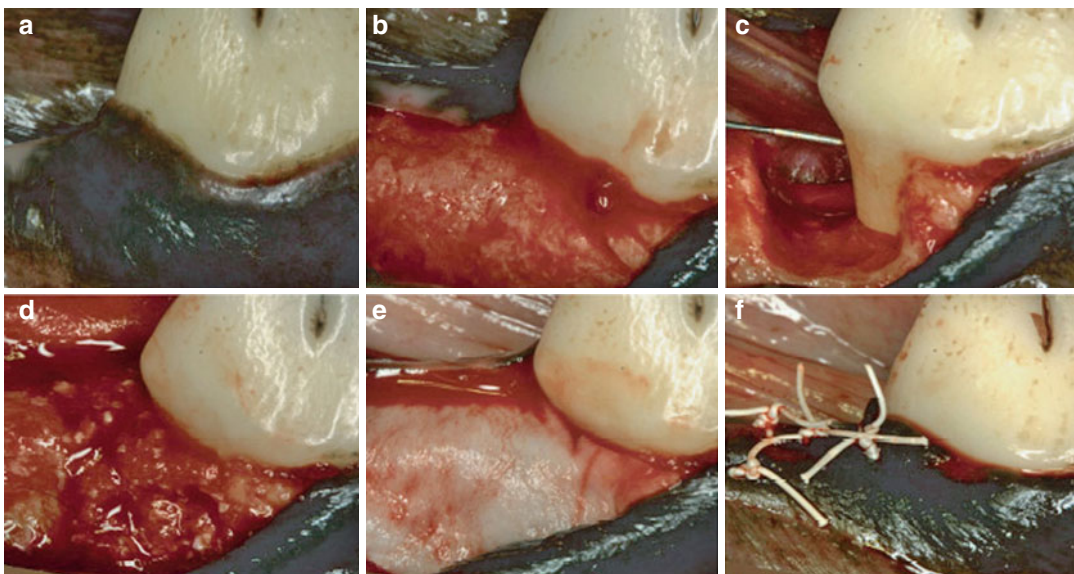


Fig. 21.26 Pre-surgery view of the experimental area (a), after flap elevation (b), and creation of a box-type, narrow/shallow (4×4 mm) intrabony periodontal defect at the mesial aspect of the mandibular first molar (c). The defect was filled with a piece of a bovine bone/collagen

matrix (d), covered with a porcine Type I collagen membrane (e) and the flaps were repositioned and sutured for primary intention healing (f) (Stavropoulos and Wikesjö 2010. Reprinted with permission from John Wiley & Sons, Inc.)



Fig. 21.27 Surgically created, critical-size, one-wall, intra-bony periodontal defect at the distal aspect of the second and mesial aspect of the fourth mandibular premolar teeth (*left*). Application of rhGDF-5/PLGA (*left centre*). Mucoperiosteal flaps adapted and sutured for primary

intention healing (*right centre*). Healing at 8 weeks (*right*). rhGDF-5, recombinant human growth/differentiation factor-5; PLGA, poly-lactide-co-glycolide-acid (Min et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

Following an incision and reflection of the gingival flap, a 3-mm-wide and 5-mm-high defect was prepared surgically at the interproximal area (**Fig. 21.28**). A thorough root planing was performed, and the flap was coronally positioned and sutured. The contralateral area was served as the control with no surgical intervention. After 8 weeks of healing, the animals were killed and the defect was analyzed histometrically (**Fig. 21.29**) and radiographically (**Fig. 21.30**). The interproximal periodontal defect resembled a naturally occurring defect and mimicked a clinical situation. After healing, the defect showed limited bone (0.89 ± 0.02 mm) and cementum regeneration (1.50 ± 0.48 mm). Moreover, this model does not require extraction of the adjacent teeth for defect preparation; thus,

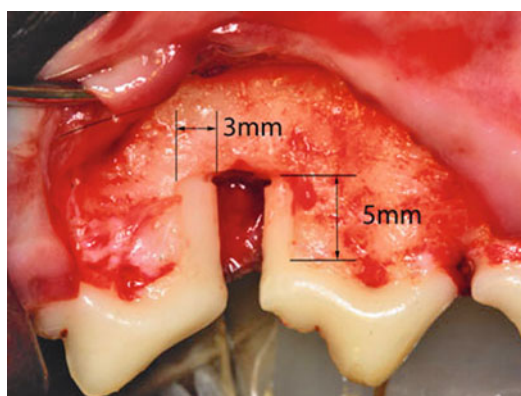


Fig. 21.28 Clinical photograph showing the configuration of the proximal no-wall defect. The alveolar bone was removed 3 mm in width from the centre of the defect to buccal and palatal bone and 5 mm in defect height (Jung et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

compared with models with intrabony defects, the experimental healing period can be reduced (Jung et al. 2011).

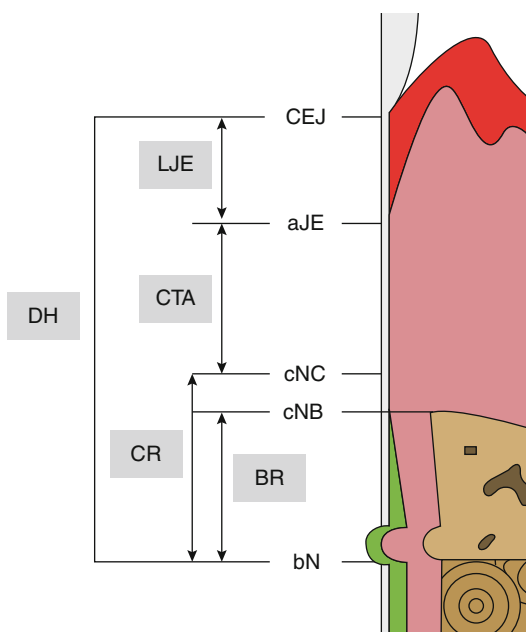


Fig. 21.29 Schematic drawing representing measurement points for histometric analysis. *CEJ* cemento-enamel junction, *aJE* apical end of a junctional epithelium, *cNC* coronal extension of new cementum, *cNB* coronal extension of new bone, *bN* base of the reference notch, *DH* defect height, *LJE* long junctional epithelial attachment, *CTA* connective tissue attachment, *CR* cementum regeneration, *BR* bone regeneration (Jung et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)

The **periodontal fenestration defect** provides a model which offers the opportunity to study regeneration of alveolar bone, root cementum and connective tissue ligament in a standardized, well-defined wound unrelated to periodontal disease. The fenestration defect has been used to investigate the healing dynamics of the periodontium and the potential of bone derivatives and engineered materials to stimulate periodontal healing (Klepp et al. 2004). The formation of critical-size fenestrations on the buccal face of maxillary canines may also make it possible to evaluate filling biomaterials or regeneration techniques (Figs. 21.31, 21.32 and 21.33) (Tal et al. 1996, 2005; Caplanis et al. 1998; Dogan et al. 2003; Klepp et al. 2004; Vastardis et al. 2005; Selvig et al. 1994; Wang et al. 1994). After full thickness flap reflection, 7mm diameter fenestrations are made through the buccal cortical plate using a trephine at slow speed under constant saline irrigation approximately 3mm apical to the crest of the alveolar bone, and the exposed root surface are thoroughly planed to remove the cementum and PDL as completely as possible (Vastardis et al. 2005).

Defects in the form of dehiscence at the level of the buccal roots of the canines, premolars or molars were also created in dogs (Fig. 21.34) (Yaffe and Shoshan 1987; da Silva Pereira et al. 2000; Kurtis et al. 2002; Sallum et al. 2004; Akizuki et al. 2005; Gineste et al. 2005a, b; Alhezaimi et al. 2009; Tomita et al. 2010).

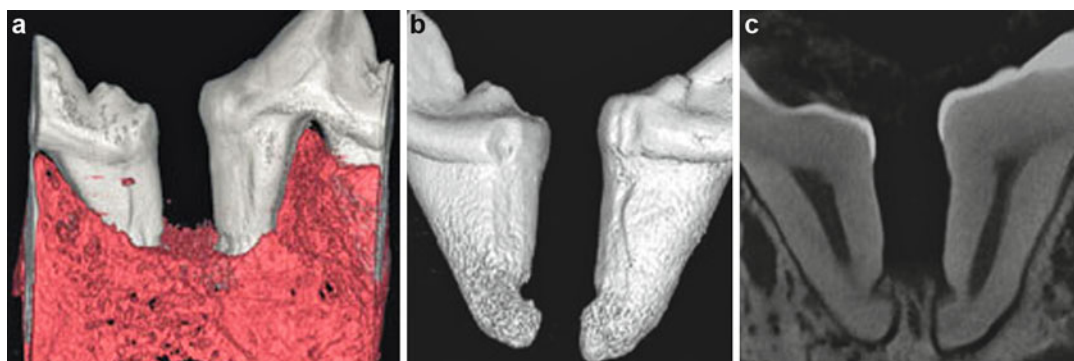


Fig. 21.30 Radiographs of specimens from the experimental group. (a) The 3D image from the micro-CT. Note the limited lateral bone regeneration from the buccal and palatal bony plate. *Pink area*: residual alveolar bone. (b) The 3D reconstructed root surface. The surface of the

defect area was smooth; the other areas were rather rough. (c) The axial cross-sectional image from the micro-CT. Interproximal bone defect remained close to the reference notch (Jung et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)



Fig. 21.31 Circular fenestration defect with peripheral and central notch (Klepp et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

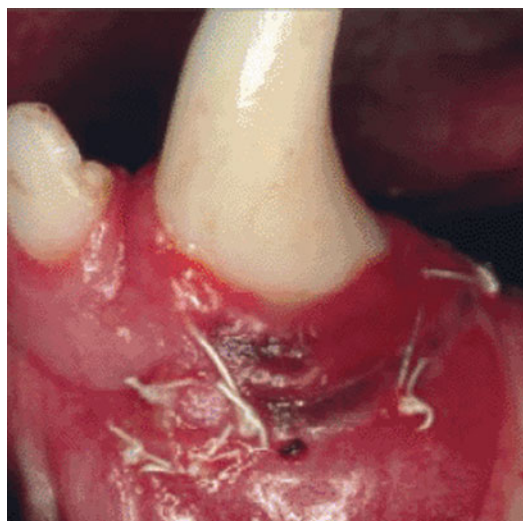


Fig. 21.33 Mandibular site following 4 weeks' healing period (Klepp et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

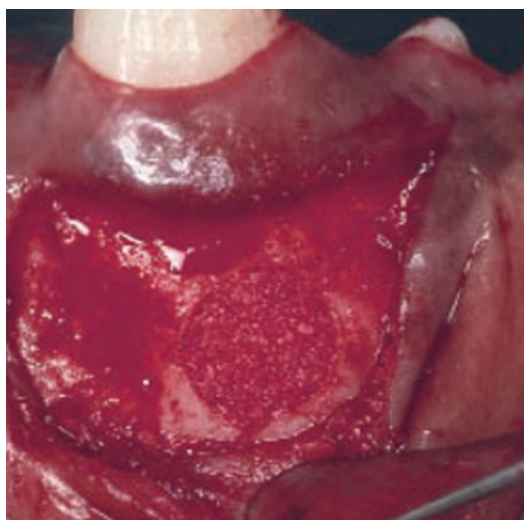


Fig. 21.32 Fenestration defect filled with demineralized freeze-dried bone allografts graft material (Klepp et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

Dogs also served as the animal model in mucogingival surgery by creating **recessions** on the canines or on second and fourth mandibular premolar which were treated either by palatine connective tissue grafts (Weng et al. 1998; Guiha et al. 2001) or by guided tissue regeneration (Casati et al. 2000; Papageorgiou et al. 2009).

Combined techniques were also described, such as:

- Surgical creation of horizontal circumferential periodontal defects (defect height from the cemento-enamel junction to the reduced alveolar crest was 4 mm) associated with placing metal strips on all exposed root surfaces for 8 weeks in aim to prevent spontaneous healing and to prevent plaque accumulation (Saito et al. 2003).
- Surgical creation of buccal osseous dehiscences by removing, with hand instruments, 4×6 mm of bone covering the mesial roots of the mandibular third and fourth premolars. In each defect, the cementum was curetted, and a strip of Toffelmire matrix band was secured to the crown of the tooth with cyanoacrylate glue, extending subgingivally into the defect area. The flaps were replaced to the original position and sutured with 4-0 Vicryl interproximal sutures. The sutures were removed 1 week later. Areas were permitted to heal without any special oral hygiene care. After 3 months of plaque accumulation, the band matrices were removed and a supragingival plaque control regimen was instituted consisting of daily brushing and topical application of 0.12% chlorhexidine digluconate for 15 days before

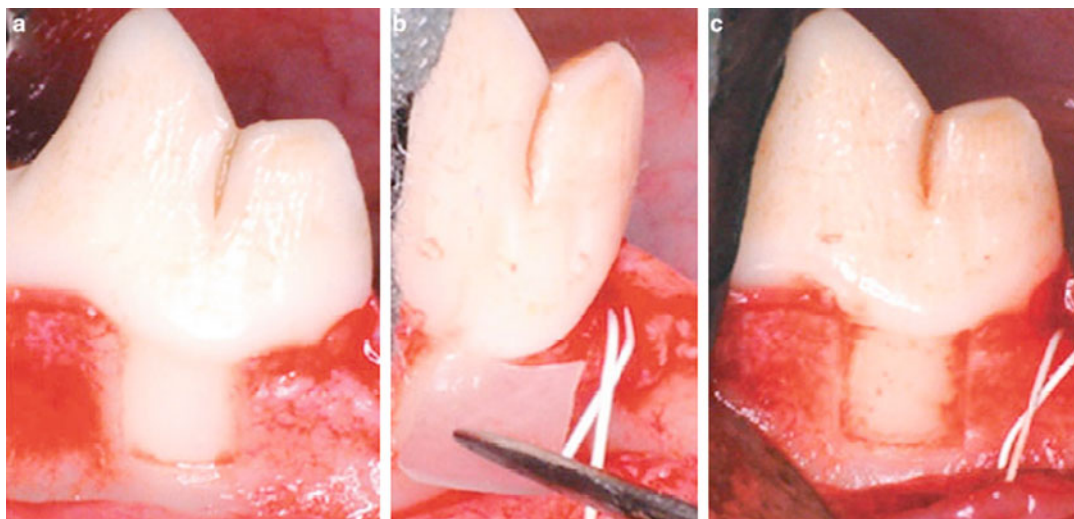


Fig. 21.34 (a) The dehiscence type of defect was formed on the buccal side of the mesial root of the mandibular first molar. Dehiscence defects were surgically prepared on the mesial roots of bilateral mandibular first molars. An intracrevicular incision was made on the buccal aspect, from mesial of the second molar to the mesial of the first molar. Following elevation of the buccal mucoperiosteal flap, a square-shaped dehiscence defect, 5 × 5 mm (width × length), was prepared on the mesial root using round

and fissure burs with sterile saline coolant (a). Root planing was performed using Gracey curettes and chisels. The cementum was completely removed. (b) Periodontal ligament cell sheet with the reinforced hyaluronic acid carrier was applied to the defect with tweezers in the experimental defect. (c) Applied periodontal ligament cell sheet with the reinforced hyaluronic acid carrier (Akizuki et al. 2005. Reprinted with permission from John Wiley & Sons, Inc.)

the surgical procedures (Pimentel et al. 2006; da Silva Pereira et al. 2000).

- Placing stainless-steel mesh (TOMY Bonding Base M, TOMY Corporation, Tokyo, Japan) (5 × 15 mm) on the denuded root cementum of maxillary canines in beagle dogs in which healthy periodontal tissue had been established by scaling. The mesh was fixed with adhesive resin cement to the crown of the canines. Subsequently, no plaque control measures were performed for 6 weeks to allow accumulation of dental plaque on the mesh. Acute inflammation was reduced after the stainless-steel mesh was removed, and a plaque control regimen, which consisted of topical application of chlorhexidine gluconate three times a week for 2 weeks, was instituted. Intracrevicular incisions were then made angle of the canine to the distal angle of the first incisor beyond the mucogingival junction. After elevation of the mucoperiosteal flap, a three-walled intrabony defect was prepared on the mesial side of the canine using round and

fissure burs with sterile saline coolant. The defect size in the maxilla was 5 × 7 × 5 mm (bucco-lingual, mesio-distal, and depth, respectively). The three-walled defects on the experimental site of the maxilla were filled with tested material, whereas nothing was applied to the contralateral site (control) (Hayashi et al. 2006).

Peri-implantitis was induced in dogs in three surgical phases that were originally described by Lindhe et al. (1992). In the first phase, extraction of the mandibular second, third and fourth premolar, and first molar (P2-M1) was performed bilaterally. After 4 months of healing, surgical implantation of screw-typed implants was performed according to a one-stage healing procedure during the second phase. Following implant insertion, a plaque control program was initiated. Tooth and implant cleaning were maintained by the use of a toothbrush once a day for 3 months. After 3 months of healing, cotton ligatures were placed in a submarginal position around each implant and the plaque control regi-

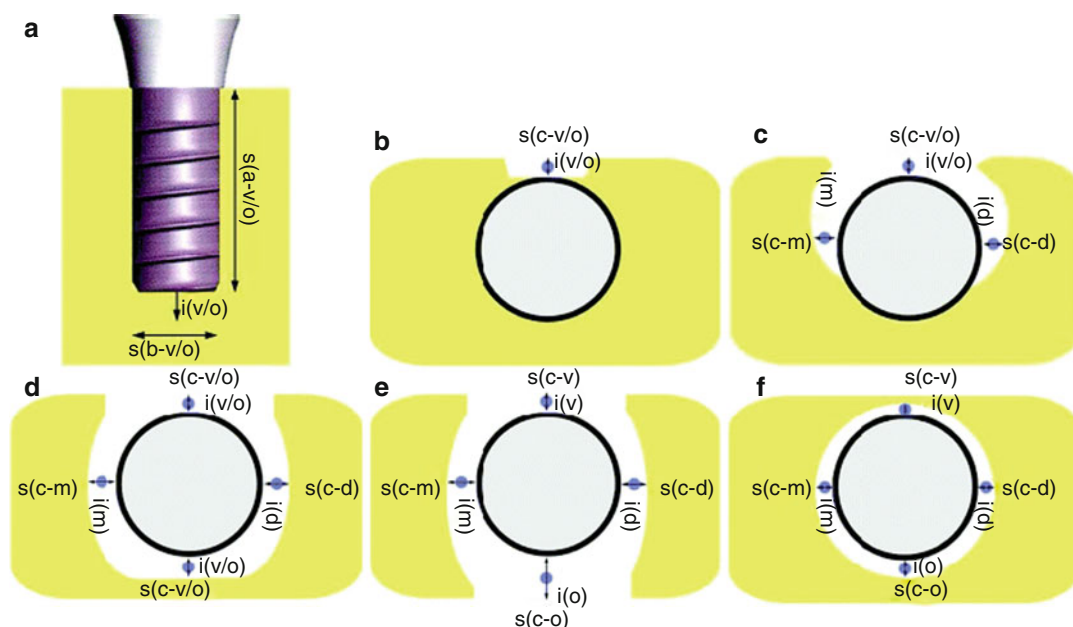


Fig. 21.35 Illustrations of Class I defects. (a) Class Ia – vestibular/oral view; (b) Class Ia – occlusal view; (c) Class Ib – occlusal view; (d) Class Ic – occlusal view; (e) Class Id – occlusal view; (f) Class Ie – occlusal view.

Arrows, *s* – component; circles, *i* – components (Schwarz et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

men was terminated. In brief, ligatures were forced into a position directly apical of the gingival margin. Subsequently, a ‘pocket’ was created that enabled the establishment of a subgingival microflora. The ligatures were exchanged once every 3 weeks and removed when approximately 30% of the initial bone support was lost based on evaluation using standardized radiomicrographs (approximately 3 months) (distance between the implant shoulder and the marginal bone level) (Schwarz et al. 2007). It has been demonstrated that configurations and sizes of ligature-induced peri-implantitis bone defects in dogs seemed to resemble naturally occurring lesions in humans (Fig. 21.35). In particular, circumferential bone loss without dehiscence of the adjacent alveolar crest was the most frequently observed type of defect configuration in both humans (55.3%) and dogs (86.6%). In these cases, circumferential bone loss was generally associated with a horizontal loss of the supporting alveolar bone (Figs. 21.36 and 21.37) (Schwarz et al. 2007).

It was demonstrated that (1) significant re-osseointegration failed to occur to implant surfaces exposed to bacterial contamination following traditional treatment of peri-implantitis lesions and (2) such integration consistently occurred at sites where a pristine implant component was placed in the bone defect following surgical debridement (Persson et al. 2001a). A treatment regimen that included (1) systemic administration of antibiotics (amoxicillin and metronidazole) combined with (2) local surface debridement and excision of granulation tissue resulted in resolution of peri-implantitis and bone fill in adjacent crater-like defects. Further, it was observed in a dog peri-implantitis model that substantial ‘re-osseointegration’ occurred to an implant with a rough (SLA) surface, while bone growth on a previously exposed smooth (turned) surface was minimal (Persson et al. 2001b). Overloading of implants in the dog model has increased bone-to-implant contact % and slightly reduced marginal bone level (Kozlovsky et al. 2007).

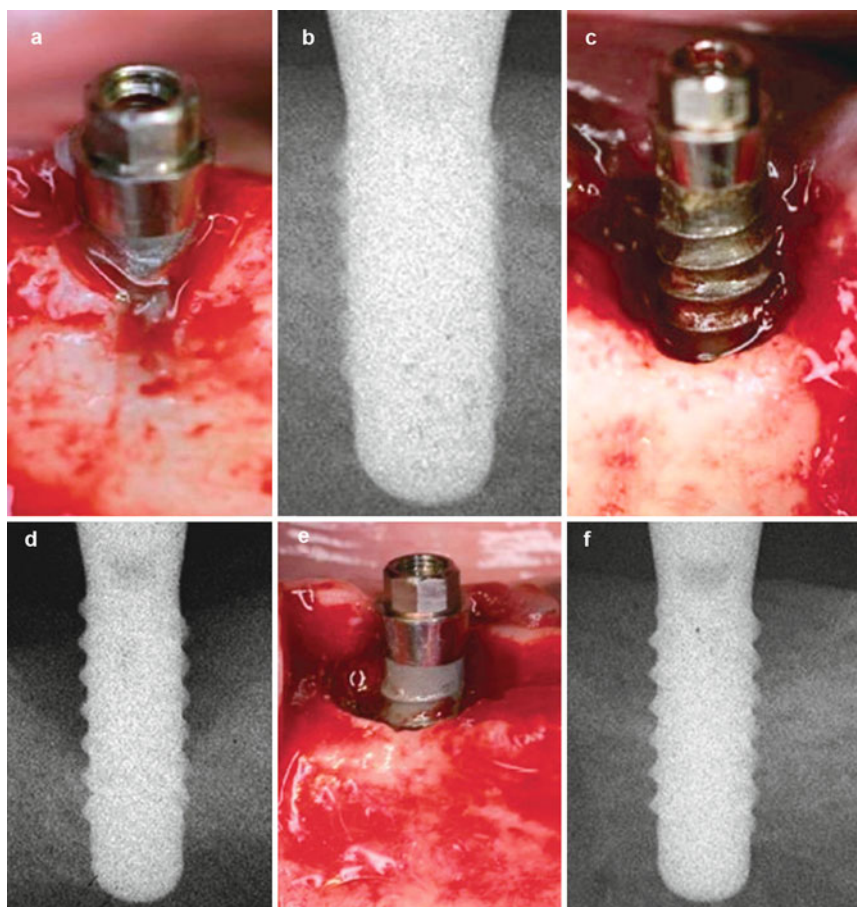


Fig. 21.36 Clinical and radiological view of ligature-induced peri-implantitis bone defects in dogs. (a, b) Class Ia defect (No. 2–034); (c, d) combined Class Ic + II defect

(No. 1–036); (e, f) combined Class Ic + II defect (No. 5–046) (Schwarz et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

21.5.2 Advantages and Disadvantages of Using Dogs in Periodontal Research

Major advantages of the dog model are the occurrence of natural disease, costs, and ease of handling. The dogs have proven similarities in response to humans, and class III furcations can be readily studied. Disadvantages include decreased numbers of class II furcations, differences from human dental anatomy, increased bone turnover rates, the limited number of defects and the more inter-animal variability (Giannobile et al. 1994).

21.6 Miniature Pigs

The oral maxillofacial region of miniature pigs is similar to that of humans in anatomy, development, physiology, pathophysiology and disease occurrence. The miniature pig has both deciduous and permanent dentitions. The initiation of tooth development and eruption in the miniature pig is quite similar to that in humans. The dental formula of the deciduous dentition for the miniature pig is Di3/3, Dc1/1, Dp1/1, Dm3/3 or Di3/3, Dc1/1, Dp2/3, Dm2/1; for the permanent dentition, it is I3, C1, P4 (3), M3 (maxillary) and I3, C1, P3 (4), M3 (mandible).

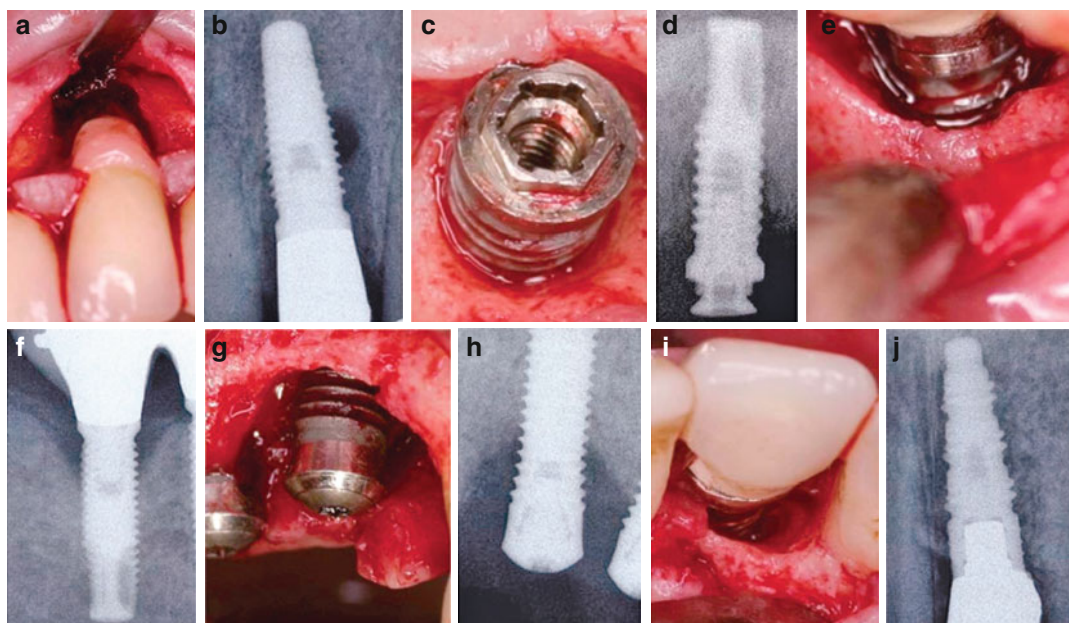


Fig. 21.37 Clinical and radiological view of naturally occurring peri-implantitis bone defects in humans. (a, b) Class Ib defect (No. 9–022); (c, d) combined Class Ie+II defect (No. 2–023); (e, f) combined Class Ie+II defect

(No. 15–047); (g, h) combined Class Ie+II defect (No. 12–024); (i, j) Class Ie defect (No. 20–021) (Schwarz et al. 2007. Reprinted with permission from John Wiley & Sons, Inc.)

The sulcular epithelium in the miniature pig consists of multilayered squamous epithelium and lamina propria of the underlying connective tissue and covers half of the crown (Meng et al. 1996). The depth of the gingival sulcus is 2–3 mm. The junctional epithelium is long and includes layers of flat cells paralleling the crown. Epithelial pegs and dermal papillae are long and slender in free gingiva and short and blunt in the sulcular epithelium. The width of attached gingiva is similar to that in humans. In the miniature pig before 6 months of age, little food debris or dental calculus can be found (Wang et al. 2007b).

Gingivitis often presents in the miniature pig after the age of 6 months. Swollen gingivae, accumulated plaque and calculus, a 1- to 2-mm red, flared zone on the gingival margin and bleeding on probing is frequently observed. There is infiltration of inflammatory cells, and blood vessels are edematous in the gingival tissues. Histologically, the inflammatory process is simi-

lar to that seen in human periodontal diseases. Serious periodontitis can be found in the miniature pig after the age of 16 months. The periodontal pocket depth (PD) may be up to 4–5 mm. Crestal resorption of the alveolar bone begins with the destruction of supporting bones (Wang et al. 2007b).

The mini pig has been successfully used to study a number of dental applications, including bone grafts, implants, enamel matrix derivative, bone regeneration at implant host sites, wound healing and periodontal tissue engineering (Wirthlin et al. 1980; Buser et al. 1998; Schliephake and Aleyt 1998; Zechner et al. 2003; Nkenke et al. 2003, 2005; Craig et al. 2004, 2006; Jensen et al. 2005, 2006; Germanier et al. 2006; Liu et al. 2008; Abukawa et al. 2009; Powell et al. 2009; Thoma et al. 2009; Pieri et al. 2009; Ding et al. 2010; Hasturk et al. 2011). **The critical-size defect (CSD) of bone mini pig mandible** with and without the periosteum was recently described by Ma et al. (2009).

Experimental periodontitis has been induced in mini pigs with different methods:

- Using ligatures, in association with bacteria, including *Porphyromonas gingivalis* (Pg), *Streptococcus mutans* (Sm) and *Actinobacillus actinomycetemcomitans* (Aa), can induce periodontitis in the Chinese miniature pig (Meng et al. 1996; Wang et al. 2007b). At weeks 4 and 8, there are significant increases in mean probing depth (PD) values and clinical attachment loss (CAL). After this active phase, there is no further destruction up to 20 weeks. Following ligature removal, PDs decreased in 3 weeks, but never returned to the original baseline values (Meng et al. 1996; Wang et al. 2007b).
- Orthodontic elastics were fixed to the premolars and molars (second and fourth quadrants) and removed again after 62 days. Additionally, metal wires were bonded to the buccal aspects of the teeth with their apical ends located at the entrance of the furcation. Before further treatment, there was a 48-day interval to allow for stabilization of the periodontal situation (Lang et al. 1998).
- The periodontal lesion was generated in the first molars area of miniature pigs by the surgical removal of bone and subsequent silk ligament suture around the cervical portion of the tooth. The created alveolar bone defect was 3 mm in width, 7 mm in length and 5 mm in depth, and notch-shaped marks were made on the root surface at the level of the top of the alveolar crest and the floor of defect (Fig. 21.38) (Liu et al. 2008).

A similar method had been applied by Hickey et al. (1991) who evaluated the microbiologic changes associated with induction of **peri-implantitis** in the micro-swine. The experimental implant abutments were ligated with 4-0 silk suture material to induce peri-implantitis for a

period of 45 days. A greater attachment level, increased probing depth, and higher gingival index and plaque index scores were noted among the experimental implants. Radiographs revealed that experimental implants showed an increased amount of bone loss when compared to control. Microbiologic studies revealed that there was a shift from Gram-positive facultative organisms to Gram-negative obligate anaerobes, including black-pigmented *Bacteroides*, in experimental implants. Singh et al. (1993) also induced peri-implantitis by the use of ligatures and a soft diet in the miniature pig.

Root surface Miller's class I gingival recessions can be surgically created in mini pigs in aim, for example, to histologically assess the interface between the acellular dermal matrix (ADM) allografts and to compare this outcome with that obtained when a connective tissue graft (CTG) is placed under the flap in control sites. Dehiscence defects were surgically created in PI, II or III (based on the presence of a minimum width of 3 mm of keratinized gingiva) at one buccal root surface in each of the four quadrants. The surgical design included two vertical incisions along the mesio-buccal and disto-buccal line angles of each root (5 mm apart and extending 7 mm apically) and the rise of a mucoperiosteal flap. The alveolar bone covering the buccal root surfaces was then removed with chisels, until creating a defect between the crestal bone and the CEJ of at least 6 mm. The flaps were then apically positioned and sutured, leaving the root surfaces exposed. The dehiscence defects created were approximately 4 mm in depth and 5 mm in width, with a remaining apical band of 3 mm of keratinized gingiva (Fig. 21.39). This surgical procedure was carried out approximately 1 month after the surgical creation of the experimental defects (Figs. 21.40 and 21.41).

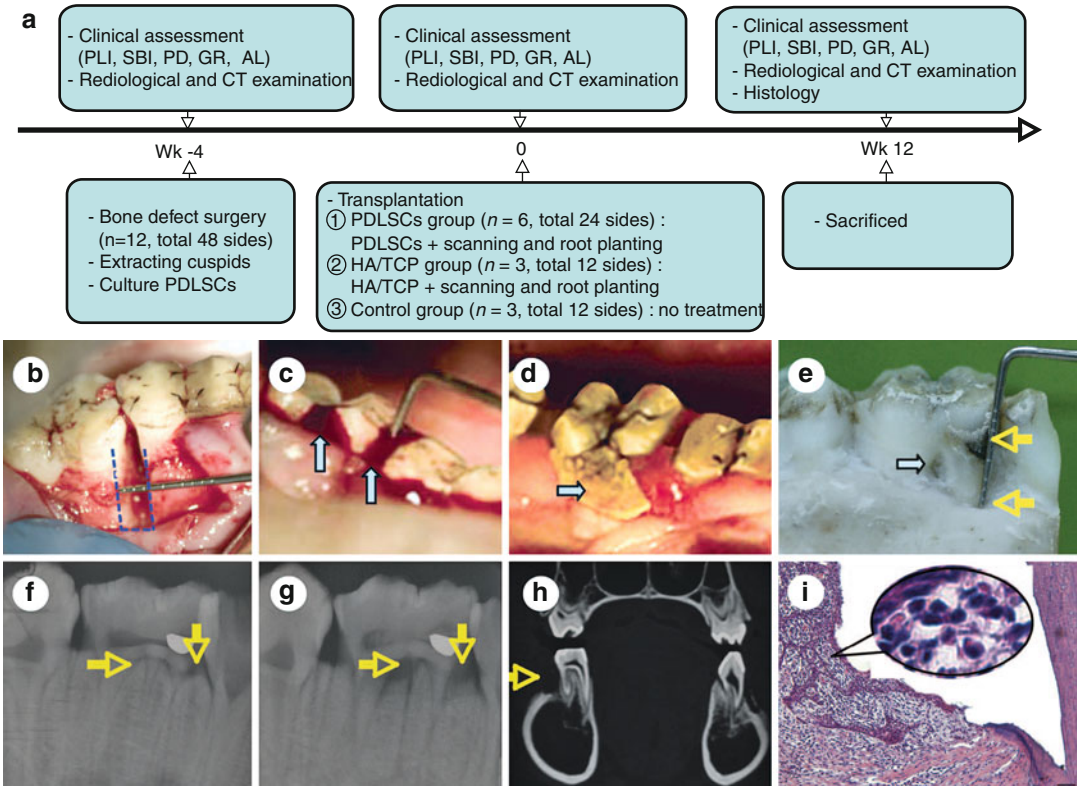


Fig. 21.38 Generation of periodontitis model in miniature pigs. **(a)** Schematic illustration of the timeline of the procedures conducted for this study. Three separate cohorts of miniature pigs were tested in this manner. **(b)** A 3-mm-wide, 7-mm-long, and 5-mm-deep bone defect was created in the mesial region of the maxilla or mandibular first molar (dotted line, arrow). **(c)** Four weeks after creating the bone defect by surgery and subsequent silk ligation, the inflammation of periodontal tissues was observed (arrows). **(d, e)** Typical periodontitis clinical findings presented in this model. Twelve weeks after the surgery, the gingiva was red and swollen. A calculus could be seen around the margin of gingiva (**d**, arrow), the inflammation was extended to the furcation area, and the whole mesial-buccal root of the first permanent molar was exposed (**e**, arrows).

(f–h) Imaging manifestations of this periodontitis model. X-ray image showed that the mineral density of the alveolar bone was decreased at the regions of furcation and the mesial side of the first permanent molar (**g**, arrows) compared with control density of the alveolar bone (**f**, arrows). CT coronal image showed obvious bone defect in the buccal alveolar region (**h**, arrow). **(i)** Histological photomicrography showing typical periodontitis histopathological manifestations, including marked periodontal tissue reduction and inflammatory infiltration. Scale bar=100 μ m. AL attachment loss, CT computed tomography, GR gingival recession, HA/TCP hydroxyapatite/tricalcium phosphate, PD probing depth, PDLSC periodontal ligament stem cell, SBI sulcus bleeding index, Wk week (Liu et al. 2008. Reprinted with permission from John Wiley & Sons, Inc.)

Fig. 21.39 The gingival recession was created surgically on the buccal aspect of the root surface having 4 mm in depth, and 5 mm in width and 3 mm of keratinized gingiva apically (Núñez et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)





Fig. 21.40 After 1 month of the experimental injury the gingival recession appeared over the buccal root surface (Núñez et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

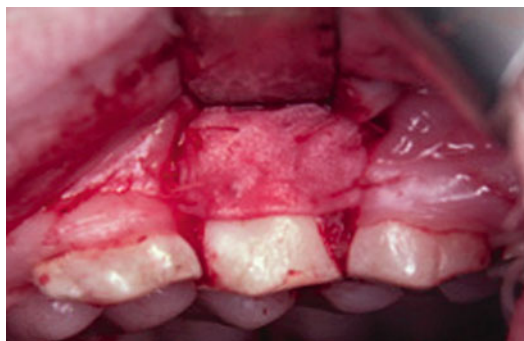


Fig. 21.41 An acellular dermal matrix (ADM) was placed in the test side and it was prepared following the manufacturer's guidelines and trimmed to fit the area. Each matrix specimen approximately 1 mm thick, was oriented with the basement membrane adjacent to the root surface and sutured into place with 5.0 polyglactin (Vicryl®) sutures (Núñez et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

21.7 Rats

Rats have one set of teeth consisting of one incisor that is rootless and three molars in each quadrant. Most histological features of the epithelium and connective tissue in the rat are similar to humans except of the sulcular epithelium, which is, keratinized. It has been suggested that this keratinization of the oral sulcular epithelium might create a better barrier to bacterial products. However, Romanowski et al. [1988] clearly showed that tracers placed in the sulcular region

enters the junctional epithelium rather than the adjacent keratinized tissues and passes preferentially along the basal lamina or within intercellular spaces closely adjacent to the tooth surface.

There is rapid wear of the occlusal surfaces with continuous eruption of the teeth and apposition of cementum and bone. This causes progressive changes in tooth position, especially the molars, which continuously move in an occlusal-distal-buccal direction (Weinberg and Bral 1999). The bony structure of the rat maxilla is altered throughout its relatively short life spans by continuous growth and the eruption and drifting of teeth. One of the more obvious changes of alveolar bone structure is a pronounced narrowing of the interdental bony septum with ageing. This loss of interdental bone width begins early in the life span of the rat, and it is believed to result from a difference in the drift rates of adjacent molars. Rat molars are known to drift distally through alveolar bone and to stimulate both resorptive activity (in advance of each drifting root) and depository activity (behind each drifting root) (Hardt et al. 1988).

The most frequently used rat strain in periodontitis studies is the Sprague–Dawley strain, but other strains have also been used successfully (Klausen 1991).

An important feature when using rodent models in periodontitis studies is to ensure that the animals do not have periodontitis at the beginning of the study. Up to 30% of regularly bred rats can develop periodontitis because of impaction of hair, food and bedding material. The use of a powder diet and wire net cages without bedding material helps to achieve periodontitis-free rats. It also should be kept in mind that the anatomic structures in the periodontium of rodents are much smaller than those in humans. Thus, because of the limited space, even an extremely small amount of bacteria in the gingival sulcus is likely to elicit a host response in rodents (Pontes Andersen et al. 2007).

21.7.1 Methods of Experimental Periodontitis Induction in Rats

Periodontal disease has been demonstrated to be induced in rats in relation to:

- (a) **Indigenous plaque** (Crawford et al. 1978; Garant and Cho 1979a, b; Heijl et al. 1980; Wyss and Guggenheim 1984; Lekic et al. 1989; Klausen 1991)

The development of gingivitis and periodontitis lesions in rats fed a normally laboratory chow diet was compared to that which occurred in rats fed a high-sucrose diet (Keyes' diet 2000). A greater incidence of periodontitis lesions with active bone resorption was noted in the diet 2000 (fortified with vitamins and sterilized by gamma irradiation) group, and at 7 weeks, there was a statistically significant difference in the size of the inflammatory cell infiltrate (Garant and Cho 1979a, b).

Investigated the microbial ecology of adherent plaque in relation to the pathological findings of gingivitis in plaque-susceptible rats. Plaque developed in the gingiva of the lower incisor in plaque-susceptible rats, but not in plaque-resistant rats, after they were fed a commercial powder diet. With increase in plaque volume, the total counts of bacteria increased 10(9)–10(11)/g. Acute gingivitis was observed within 3 months, and subacute-chronic gingivitis was observed between 2 and 12 months, suggesting that proportional changes in the gingival plaque flora may uniquely contribute to the development of gingival inflammation in rat experimental model.

However, even if sucrose was added in the diet (Heijl et al. 1980; Crawford et al. 1978; Wyss and Guggenheim 1984) or in the drinking water (Lekic et al. 1989), it was also reported that a variety of microorganisms (*Streptococcus sobrinus*, *Actinomyces viscosus*, *Porphyromonas gingivalis*) can be established in rats maintained an ordinary food (Klausen 1991).

- (b) **Experimentally introduced organisms, experimentally introduced bacterial products**

Several organisms were experimentally introduced in rats in aim to induce periodontitis: *Actinomyces naeslundii* (Strain I), *Actinomyces viscosus* (Strain T14), *Streptococcus mutans* (6715), *Actinobacillus actinomycetemcomitans*, *Actinomyces israelii*, *Porphyromonas gingivalis* (Irving et al. 1974, 1978, 1979; Heijl et al. 1980; Behbehani et al. 1983; Chang et al. 1988, 1994; Lekic et al. 1989; Klausen et al.

1989b; Suzumura et al. 1989; Klausen 1991; Katz et al. 1996, 1999; Doxey et al. 1998; Karimbux et al. 1998) (Socransky et al. 1970)

Various methods for introducing suspected periodontal pathogens in rats were used:

- Expose the rats to mono-contaminated faces or mono-contaminated rats (Irving et al. 1974, 1978; Heijl et al. 1980).
- Intraperitoneal injections with 10⁹ formalin-killed cells of microorganisms, two times, when the rats were 4 and 6 weeks old (Lekic et al. 1989).
- Chang et al. (1988) applied the test organisms (*Streptococcus mitis* (strain KM), *Actinomyces odontolyticus* (Strain GM), *Capnocytophaga ochracea*) to the oral cavity and gingiva as follows: an 18–48-h anaerobic culture in pre-reduced P.Y. broth was prepared, transferred to germ-free isolators, and brushed onto the dentition of the germ-free rats using a sterile #8 camel-hair brush. Approximately 0.25 ml of the suspension was applied by inducing the animal to bite into the brush to aid the inoculation of the gingival crevice. This procedure was repeated daily for up to 5 days until the test microorganisms were positively identified after periodically sampling the subgingival plaque from each animal.
- A different technique was used for the strict anaerobic requirement of *P. gingivalis* (Chang et al. 1994; Klausen 1991; Katz et al. 1996, 1999). Freshly harvested *P. gingivalis* 33277 (approximately 10¹⁰ cells/ml) was mixed with 2% carboxymethylcellulose, and individual rats were given 0.5 ml of the suspension on day 13 following immunization by gastric intubation with the aid of an intubation needle. This procedure was repeated twice at 24-h intervals (Katz et al. 1999), at 48-h interval (Nakajima et al. 2006; Okada et al. 2010) or once a day for a 3-day period (Chang et al. 1994; Doxey et al. 1998; Karimbux et al. 1998).

Behbehani et al. (1983) reported that the implantation of *Actinomyces israelii* in gnotobiotic rats necessitated repeated inoculation before the bacteria appeared to establish themselves in the oral cavity. The delayed implantation of *A.*

israelii in germ-free rats observed in the present experiments suggests that this organism may be deficient in some aspect which affects the initial events leading to colonization of the mouth. Thus, *A. israelii* appears to be different in this respect from *A. viscosus* and *A. naeslundii*, which readily infect gnotobiotic rats. Gross evidence of colonization of *A. israelii* was first noted as small colonies associated with impacted hair shafts and particles of bedding material, particularly between upper first and second molars. Periodontal pathology was initiated in these areas and was accompanied by accumulation of high numbers of PMNs. Once established in these interproximal areas, *A. israelii* was able to proliferate into large plaque-like masses. However, these plaque samples were more subgingival and of a harder consistency than plaque formed by *A. viscosus* and *A. naeslundii* (Behbehani et al. 1983).

Klausen et al. (1989b) had used for experimental periodontitis inoculation *A. viscosus* ATCC 19246, *P. gingivalis* ATCC 33277 and oral Spirochetes strain with two endoflagella from each cell end. One millilitre of each strain was inoculated peroral administration daily for three consecutive days. The spirochete strain was re-

inoculated by similar procedures 8 weeks after the first inoculation. Both 2 weeks and 90 days after the inoculation, *A. viscosus* and *P. gingivalis* were routinely recovered from oral swabs and dental plaque samples, whereas spirochetes could not be detected at any time.

Mono-bacterial (single pathogen) periodontal infections primarily using *P. gingivalis*, *T. denticola*, *T. forsythia*, or *F. nucleatum* have been studied in rats and mice (Baker et al. 1999; Lalla et al. 2003; Sharma et al. 2005; Kesavalu et al. 2007). While these studies have generally used mono-infections in animal models, increasing evidence supports the concept that bacterial interactions among selected members of the subgingival microbiota at any one time during periodontal disease progression are important (Verma et al. 2010). Verma et al. (2010) investigated mixed microbial (two species) periodontal disease using *T. forsythia* and *P. gingivalis* and examined their colonization/infection characteristics, induced inflammation, immune response profiles, cytokine/chemokine responses and induction of alveolar bone resorption. It was hypothesized that *T. forsythia* and *P. gingivalis* demonstrate synergistic virulence (colonization/infection) for

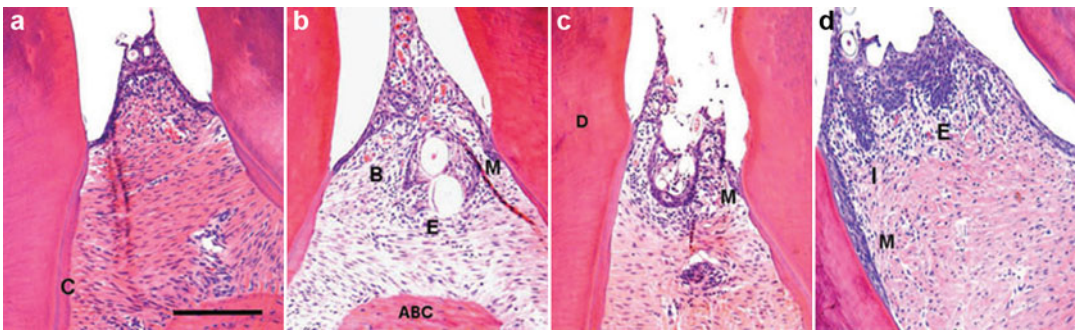


Fig. 21.42 Histopathologic evaluation of inflammation. Comparative maxillary histology (hematoxylin and eosin staining) of alveolar bone sections from the maxilla of rat infected with *P. gingivalis* and/or *T. forsythia* at 12 weeks. (a) Section from a control uninfected rat displaying dentin (D) and cementum (C) with minimal inflammation and hyperplasia of the crevicular epithelium. (b) Section from the *P. gingivalis* infected rat displaying significant inflammation (I) and hyperplasia of the epithelium with elongation of the rete ridges (E) and increase in small capillary infiltration (B). Apical migration of JE (M) is noted in the section. Alveolar

bone crest (ABC) is seen in some of the sections. (c) Section from the *T. forsythia* infected rat also displaying features very similar to image (b). Apical migration of JE is also noted. (d) Section from the mixed infection with *P. gingivalis*+*T. forsythia* exhibiting most severe histological changes namely, significant inflammation (I), hyperplasia of the epithelium with elongation of rete ridges (E), significant apical migration of the JE (M) and numerous dilated capillaries. All images in the panel are at a magnification of 10 $\times$. Scale bar represents 100 μ m (Verma et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

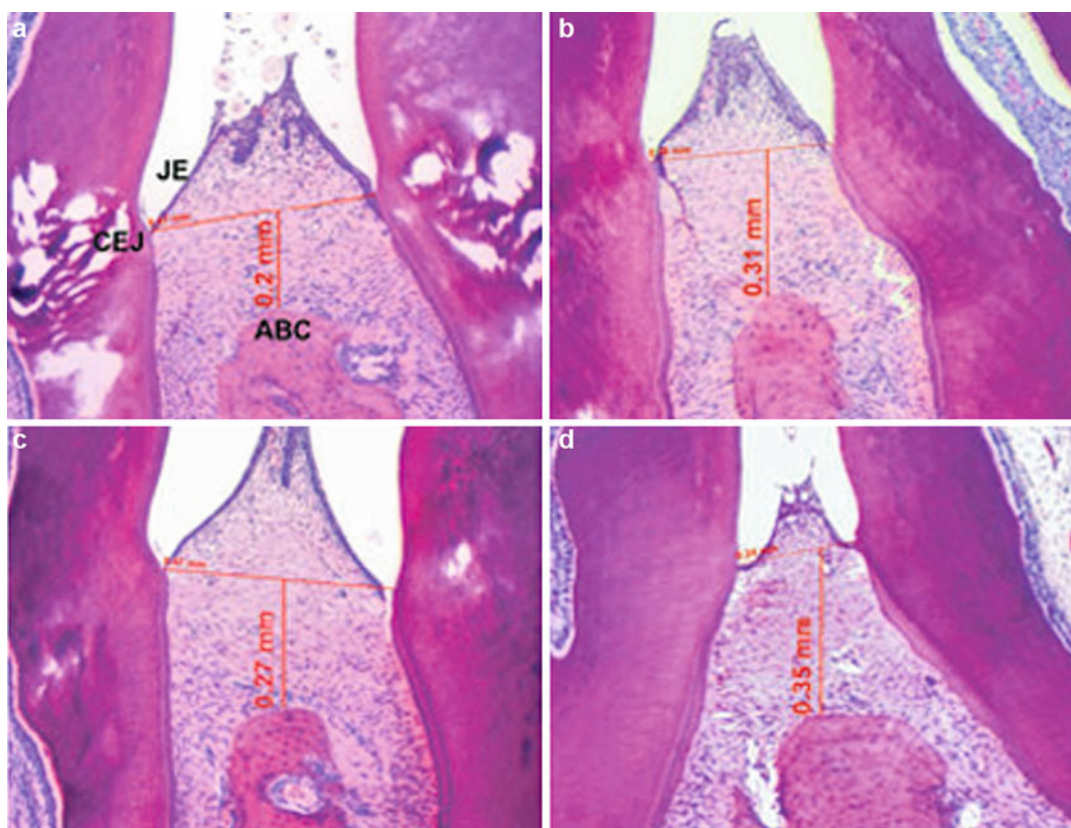


Fig. 21.43 Histometric analysis of bone loss. Photomicrographs illustrating the top of the alveolar bone of the maxilla showing the interdental alveolar bone distance by histometry (hematoxylin and eosin staining). The distances between the ABC and the CEJ were measured in mm using the AxioVision LE 29A software version 4.6.3. Image analysis program. Representative image of a control

rat showing the distance between the ABC and CEJ (a). (b) Rat right maxilla infected with *P. gingivalis*, (c) *T. forsythia*, and (d) *P. gingivalis*+*T. forsythia* during 12 weeks of infection showing increased distances between the top of the alveolar bone and the CEJ intersection. JE junctional epithelium (Verma et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

inducing chronic periodontal inflammation leading to enhanced alveolar bone resorption (Figs. 21.42 and 21.43) (Verma et al. 2010). The amount of each bacterium, *Porphyromonas gingivalis* strain 381 and *T. forsythia* ATCC 43037, was determined quantitatively, and the organism was resuspended at 2×10^{10} bacteria/ml. For mono-infection, each bacterium (2×10^{10} cells/ml) was mixed with equal volumes of sterile 2% (w/v) low-viscosity carboxymethylcellulose (Baker et al. 1999; Lalla et al. 2003; Lee et al. 2006; Kesavalu et al. 2007) with PBS and 0.5 ml was used for infection (5×10^9 cells) by oral gavage as described previously (Baker et al. 1999;

Rajapakse et al. 2002; Lalla et al. 2003; Sharma et al. 2005; Lee et al. 2006; Kesavalu et al. 2007). For oral mixed microbial infection, *T. forsythia* (2×10^{10} cells/ml) was gently mixed with an equal volume of *P. gingivalis* (2×10^{10} cells/ml) and allowed to interact for 5 min. An equal volume of sterile 2% CMC was added, mixed thoroughly, and 0.5 ml (2.5×10^9 cells of *T. forsythia*, 2.5×10^9 cells of *P. gingivalis*) was administered by oral gavage (Verma et al. 2010). The synergistic effects of *Porphyromonas gingivalis*, *Treponema denticola* and *Tannerella forsythia* in experimental periodontitis in rats were also demonstrated (Kesavalu et al. 2007).

Suzumura et al. (1989) had investigated, histologically and histometrically, the effect of bacterial protease on the periodontal tissues in rats. 0.03-ml bacterial protease (distilled water with 300 units, Sigma type XIV) was topically applied to the lingual marginal gingiva of the mandibular incisor by placing a small amount of cotton containing that solution on the gingiva for 1 min. The animals of both groups were treated once a day for 1, 3, 6, 9, 15 and 21 days, and they were killed 1 days after the final treatment. A marked apical proliferation of the junctional epithelium and a slight inflammatory cell infiltration subepithelially was observed in the experimental group at 9, 15 and 21 days.

Extractable lipids from *Actinomyces naeslundii*, strain I and *Streptococcus mutans* (6715) were injected into the maxillary gingiva lingual to the mesial cusp of the first molar tooth. The animals were sacrificed at intervals, 6 h to 14 days, and the tissues examined histologically. The lipids from *A. naeslundii* gave a tissue response different than those from *Streptococcus mutans*; however, lipids from both bacterial sources caused bone changes (Irving et al. 1979).

The lipopolysaccharide (LPS) component of the cell wall of Gram-negative bacteria is a significant inflammatory stimulus that triggers an innate immune response. Thus, the injection of LPS into the gingival tissues is a model for examining how the innate immune response to this bacterial component induces inflammation to stimulate osteoclastogenesis and bone loss. This model produces a histopathological aspect similar to the other models and to that observed in established periodontitis in humans, characterized by increased infiltration of leukocytes, higher levels of proinflammatory cytokines, collagen degradation and alveolar bone resorption (Ijuhin et al. 1992; Miyauchi et al. 1998). Alveolar bone loss has been observed as early as 7 days after the start of injections of LPS from various microorganisms, including *Escherichia coli*, *A. actinomycetemcomitans* and *Salmonella typhimurium*, suggesting that the source of LPS may not significantly affect this particular outcome (Graves et al. 2012). Typically, defined

amount of purified bacterial LPS suspended into small micro-volumes (1–6 µl) is injected into the gingival tissues surrounding the posterior teeth of either mice or rats. The technique is somewhat sensitive and requires the use of micro-syringes for small volumes and thin needles (Graves et al. 2012). For example, in the experiments performed by Vardar-Sengul et al. (2008) and Ramamurthy et al. (2002a), the rats were injected with 10 µl saline or *Escherichia coli* endotoxin (serotype 055:B5, L2637; 1 mg/ml) into the palatal gingiva between the first and second maxillary molars and into the labial and oral gingiva between the central incisors in the mandible and maxilla every other day for 3 days. All rats were anaesthetized and sacrificed on day 15 (Vardar-Sengul et al. 2008; Ramamurthy et al. 2002a). Kirkwood et al. (2007) delivered 2 µl of a 10 mg/ml solution of *A. actinomycetemcomitans* LPS to the palatal gingiva via a 33-gauge Hamilton syringe between the maxillary first and second molars three times per week for 8 weeks (480 µg of LPS over the 8-week period). After injection of LPS into rat or mouse gingiva, both bone loss and connective tissue breakdown have been reported (Kawai et al. 2000; Llanvaneras et al. 2001; Ramamurthy et al. (2002a, b); Kirkwood et al. 2007; Vardar-Sengul et al. 2008; de Aquino et al. 2009). The time course of inflammation and bone loss following a single injection of *Salmonella typhimurium* in the rat was also evaluated (Dumitrescu et al. 2004). Periodontitis was induced by intragingival injection of 1 µl of LPS (10 µg/µl) derived from *Salmonella typhimurium* in sterile saline. A fine hypodermic needle was inserted at the mesio-lateral aspect of the first right mandibular molar, and the tip moved distally so that the injection was made at the interdental papilla between the first and second molars. The total volume of injection was reduced from 10 µl as originally described to 1 ml, as in initial experiments, the tightly bound tissue of the gingivae would not expand to accommodate an injection volume of 10 ml, and much of the injected volume was lost along the needle track. The injection was made slowly, and the needle held in place for several seconds post injection to ensure that the LPS was not lost

through the needle track. LPS injection resulted in a significant gingival and periodontal inflammation with inflammatory infiltrate, apical migration of the junctional epithelium, interdental bone loss and activation of osteoclasts at the site of injection 7 and 10 days after injection. At 10 days post injection, there was a significant trend for bone loss on both sides of the mandible.

Topical application of lipopolysaccharide (*Escherichia coli* LPS) and proteases (*Streptomyces griseus* proteases) for 4 or 8 weeks induced not only periodontal inflammation but also an elevation in the serum lipopolysaccharide concentration, with increasing hepatic inflammation, steatosis, and 8-hydroxydeoxyguanosine levels in a time-dependent manner

(Fig. 21.44). 25 $\mu\text{g}/\mu\text{l}$ LPS (0.5 $\mu\text{l} \times 3$ times) and 2.25 U/ μl proteases (0.5 $\mu\text{l} \times 3$ times) were introduced by micropipette into the gingival sulcus of both maxillary first molars. There was an interval of 10 min between each application of LPS or proteases (Ekuni et al. 2006; Tomofuji et al. 2007, 2009b; Ekuni et al. 2008).

(c) **By placing a ligature** (Sallay et al. 1982; Shoji et al. 1995; Lohinai et al. 1998; Breivik et al. 2000a, Bezerra et al. 2000; Gašperšič et al. 2002; Tomofuji et al. 2008, 2009a; Busch et al. 2008; Toker et al. 2008; Ekuni et al. 2009; Huang et al. 2010)

The placement of a ligature (sterilized black braided 000 silk ligature, non-absorbable surgical suture A-54 Ethicon, Ethicon INC, Somerville, NJ) around the second maxillary

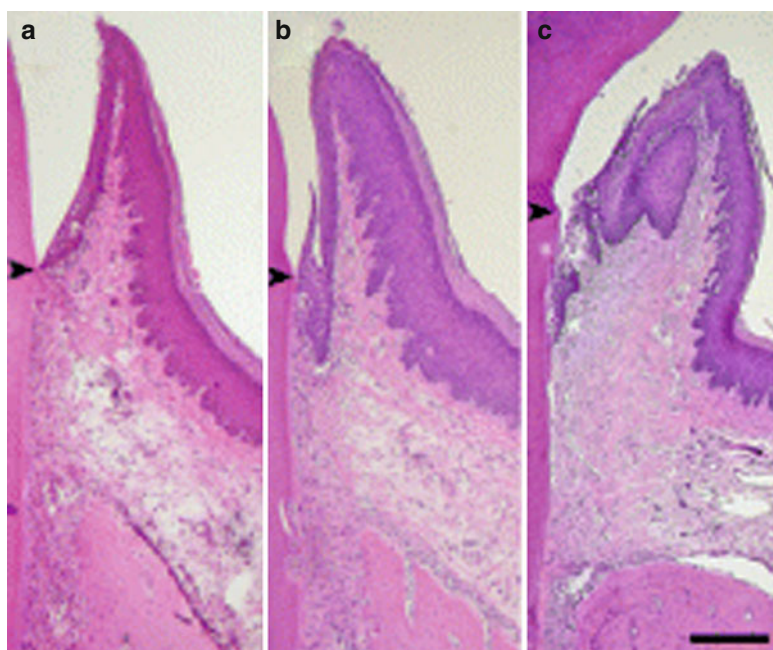


Fig. 21.44 Induction of histological changes consistent with periodontitis. Lipopolysaccharide with protease was applied daily to the palatal gingival sulcus of maxillary molars and animals were killed by perfusion fixation at time 0 (a), and at 2 weeks thereafter. The jaws were removed, decalcified, and 8 weeks (c) (b) paraffin embedded, sectioned and stained with hematoxylin and eosin. Representative sections are shown, and the cemento-enamel junction (CEJ) Mean increase in apical epithelial is designated with black arrows. (d) Cell migration past

the CEJ and along the root surface was measured in all groups using a microgrid, and the mean alveolar bone to CEJ distance was calculated. Mean bone loss at 2 and 8 weeks represented the subtracted differences over the time 0 controls. All values are shown as mean $\pm$ standard error (SE) ($n=7$). Apical epithelial migration and bone loss was significantly induced at 8 weeks ($P<0.001$). Scale bar (black), 200 μm (Ekuni et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

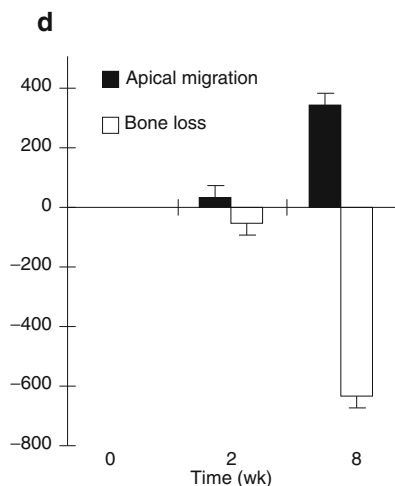


Fig. 21.44 (continued)

molar (Sallay et al. 1982; Breivik et al. 2000a, Bezerra et al. 2000a) and knotted on the vestibular side (so that it remained subgingivally in the palatal side and supragingivally in the vestibular side) or around the cervix of the lower first molar and knotted mesially (Lohinai et al. 1998), in the conventional rat, caused extensive alveolar bone resorption at 9 and 14 days after placing the ligature. The ligature also caused leukocytosis in the peripheral blood (Sallay et al. 1982).

At 42 days, the ligature biofilm in rats presents various bacterial species that hybridize to probes of periodontal bacterial species commonly observed in human. Twenty-five species were detected in the ligatures. *Streptococcus*- and *Actinomyces*-like species were found in high-mean counts, followed by *Fusobacterium*-, *Prevotella nigrescens*- and *Parvimonas micra*-like species and *Porphyromonas gingivalis*- and *Aggregatibacter actinomycetemcomitans*-like species (Duarte et al. 2010).

Even though the LPS injection model has been shown, both *in vitro* and *in vivo*, to initiate chronic inflammation in the connective tissue by inducing expression of various inflammatory mediators and ultimately result in alveolar bone loss, the ligature

model of experimental periodontitis is considered by some to be more representative of periodontitis in humans mainly because of the participation of live microorganisms with diverse virulent factors other than LPS, known as pathogen-associated molecular patterns. This greater diversity of antigens may result in a more complex host response, which may have an effect on the cytokine and inflammatory mediators' network, including matrix metalloproteases (de Aquino et al. 2009).

The application of ligature model can only be recommended for short (<15 days) observation periods. Possibly, the age-dependent remodelling processes are factors which contribute to the diminution of the ligature effect after day 15, since the movement of the teeth away from the sulcus necessarily also distances the ligatures from their site of efficacy (Kuhr et al. 2004). De Aquino et al. (2009) reported a decrease on the severity at 30 days, which hypothetically could be ascribed to a protective feature of periodontal tissues that receded apically from the aggression located on the gingival margin, in an attempt to recover the biological space. This possibility is supported by the fact that once placed, ligatures were kept throughout the 30-day experimental period; however, they were not pushed further apically even if the gingival margin had receded (de Aquino et al. 2009). Effect of ligature is restricted to adjacent structures; systemic effect of ligature has not been described (Kuhr et al. 2004).

- (d) **Applying a standardized elastic ring** (diameter 2.0 mm, thickness 1.3 mm, width 0.9 mm) around the cervix of the right mandibular first molar (Irokawa et al. 1977; Shoji et al. 1995, 2007; Mitsuta et al. 2002;). After 8 days, the presence of an elastic ring around the neck of the mandibular first molar induced an acute inflammatory reaction in periodontal tissues: vertical resorption of the bone in the interdental area between the first and second molars, widening of the ligamental space around the roots of the first and second

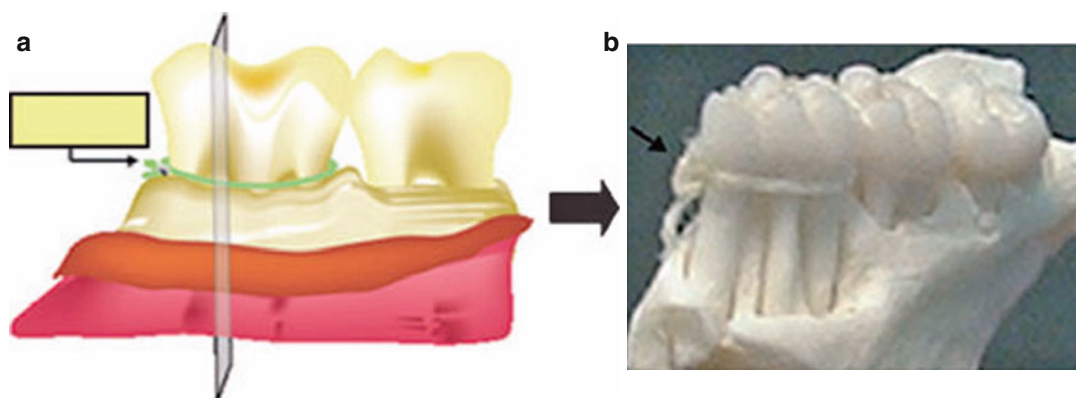


Fig. 21.45 Schema of the surgical procedure (a) and typical bone destruction (b). (a) After the reflection of full-thickness mucoperiosteal flaps were made at the right and left first molars of the mandibula, the periodontal ligament and cementum were mechanically removed. Then, silk thread (Φ 0.1–0.15 mm), soaked for 1 day in Brain–Heart Infusion Broth containing 5 mg/ml of yeast extract, was ligated around the cervix of the first molars.

(b) Alveolar bone destruction in the first mandibular molar 14 days after the induction of experimental periodontitis. Significant bone resorption in the first molar was observed as compared with the second and third molars. The arrow shows the ligature in the mandibular first molar (Fujishiro et al. 2008. Reprinted with permission from John Wiley & Sons, Inc.)

molars, together with significant bifurcation involvement, were noted.

- (e) **Periodontal fenestration defects** (4 mm in width, 3 mm in length and $\approx$ 1 mm deep) can be created at the buccal aspect of the distal root of first mandibular molars, as described by Benatti et al. (2005, 2006). Briefly, the superficial bone is removed using a round dental bur (diameter 2 mm), at slow speed under saline irrigation, and the distal root of the first mandibular molar was denuded of its periodontal ligament, cementum and superficial dentin using a chisel. Neither antibiotics nor anti-inflammatory drugs were administered after surgery. The animals were killed 3 and 6 weeks postoperatively, and the mandibulae were retrieved for histological evaluation.

(f) **Mixed models**

Fujishiro et al. (2008) induced rat experimental periodontitis by elevating a full-thickness gingival flap and ligating silk threads around the first molars of the mandible. Under general anaesthesia (intramuscular injection of sodium pentobarbitone), the reflection of full-thickness mucoperiosteal flaps were made at the right and left

first molars of the mandibula. Periodontal ligament and cementum were mechanically removed. Then, silk thread (Φ 0.1–0.15 mm), soaked for 1 day in Brain–Heart Infusion Broth containing 5 mg/ml of yeast extract, was ligated around the cervix of the first molars, as shown in Fig. 21.45a, to promote the accumulation of plaque and delay the wound healing of the flap. Bone resorption continued until 14 days after the induction of experimental periodontitis, and the bone defect form of the furcation was similar to that of the class II furcation involvement (Lindhe’s classification), as shown in Fig. 21.45b. Therefore, it was suggested that this experimental periodontitis model might be suitable for investigating the effect of Emdogain. Because rapid bone resorption and the appearance of many osteoclasts and a large number of cells expressing interleukin-1 β and RANKL were observed at 7 days and maximum bone resorption was observed at 14 days after inducing periodontitis, Emdogain or propylene glycol alginate treatment was carried out at 14 days in this study (Fujishiro et al. 2008).

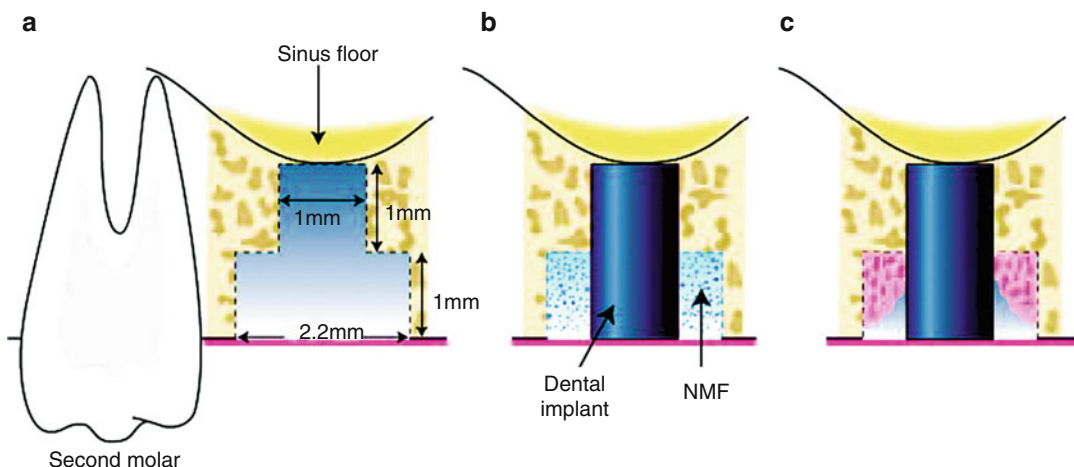


Fig. 21.46 Infrabony peri-implant defect in the rat. (a) After extraction of the maxillary first molar, the socket was allowed heal for ~4 weeks. (b) During the second surgery, a bone defect was created by means of an osteotomy in the location of the former tooth. An implant (1 × 2 mm) was press-fit into position, and the NMF was delivered to

the peri-implant bone defect, followed by soft-tissue wound closure. (c) At multiple time-points (10, 14 and 21 days), tissue samples were harvested, and bone healing was evaluated (Pellegrini et al. 2009. Reprinted with permission from Sage Publications)

- (g) Other models for oral and periodontal reconstructive therapies are (1) **alveolar extraction socket model**, (2) **infrabony peri-implant defect model** (Fig. 21.46) and (3) **vertical bone augmentation (capsule model)** in the rat (Fig. 21.47) (for review, see Pellegrini et al. 2009).

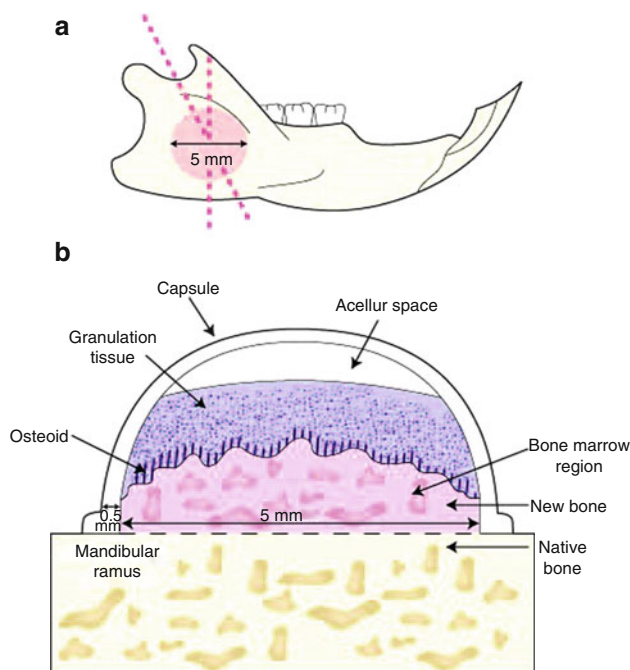
21.7.2 Clinical and Histological Findings in Experimental Periodontal Disease in Rats

The clinical and histological findings in experimental periodontal disease in rats are similar to findings in man (Klausen 1991). Gingival bleeding was observed on days 3, 7 and 14 after healthy rats were placed in contact with *Streptococcus mutans* 6715, *Actinomyces viscosus* ATCC 33277, *Fusobacterium nucleatum* ATH101 and *Bacteroides gingivalis* ATCC 19246 (Lekic et al. 1989).

Heijl et al. (1980) had demonstrated that germ-free rats of the Sprague–Dawley strain and rats mono-infected with microorganisms frequently associated with human gingivitis, periodontitis

and periodontitis lesions: *Actinomyces naeslundii*, *Eikenella corrodens*, *Bacteroides asaccharolyticus* and *Campylobacter sputigena* may suffer severe breakdown of the periodontal tissues. With increasing experimental time, the degree of alveolar bone loss increased in all groups except of germ-free controls. In germ-free rats, bone loss increased up to day 56 following which no further bone degradation could be recognized. Periodontal tissue destruction was related to the impaction of hair and bedding material. In fact, only in areas with severe impaction of foreign material could a destruction of the supporting apparatus be recognized. The tissue lesions observed in histological sections from various groups of rats had many features, which were clearly different from those characterizing gingivitis and destructive periodontitis in humans, monkeys and dogs. First of all, in no section out of more than 2,000 examined in the light microscope was it possible to identify dental plaque. Second, in the rats, periodontal tissue destruction was frequently noted in the interdental areas. In variance with periodontal disease in humans (monkeys and dogs), the central portion of the interdental space was

Fig. 21.47 Vertical bone augmentation in rat (capsule model). **(a)** The Teflon capsule was positioned on the vestibular surface of the mandibular ramus and stabilized by means of sutures. **(b)** The pattern of bone neogenesis included the deposition of loose granulation tissue, subsequently replaced by osteoid and new bone. The newly formed mineralized tissue was in close contact with the original bone of the mandibular ramus. An acellular region was noted superior to the surface of the granulation tissue. Tissue samples were harvested from 2 to 6 months, and bone neogenesis was evaluated (Pellegrini et al. 2009. Reprinted with permission from Sage Publications)



consistently more severely affected than areas adjacent to tooth surfaces. This phenomenon has been referred to as '**interdental cratering**'. In the sections obtained from rat, inflammatory cells were frequently present in a narrow zone beneath the epithelium of the 'cratering' part of the interdental papilla. These infiltrates which in cases of active bone resorption extended to the bone surface harboured neutrophils and macrophages but were consistently devoid of plasma cells, lymphocytes and blast cells (Heijl et al. 1980).

Histologically, after large adherent plaque deposits were present in interdental spaces, migration of the epithelial attachment and an increased incidence of rete peg formation can be commonly observed (Garant 1976a). Many polymorphonuclear leukocytes can be detected in the plaque and on the surface of the gingiva, having presumably migrated through the sulcus (Irving et al. 1974). An inflammatory cell infiltrate containing T and B lymphocytes, macrophages and polymorphonuclear leukocytes appears in the connective tissue (Garant 1976a; Irving et al. 1974; Lekic et al. 1989).

Different changes were induced after inoculation with Gram-negative organisms compared with Gram-positive bacteria. No plaque was found at any stage of the experiment when gnotobiotic rats were mono-infected with a Gram-negative, anaerobic rod isolated from a case of human periodontitis (Irving et al. 1975b). Inflammatory reactions, as evidenced by the infiltration of leukocytes, were minimal; nor were more lymphocytes than normal present. No plasma cells were seen, but they are normally rare in rats. The bone destruction was accompanied by numerous osteoclasts, present in considerable greater numbers than normal on the crest and other aspects of the alveolar bone (Irving et al. 1975b, 1978). After mono-infection with Gram-positive (*A. naeslundii*), large adherent plaque deposits were present in the interdental spaces. The alveolar bone of these interdental areas invariably demonstrated scalloping of the surface and several small osteoclasts. These small osteoclasts frequently contained no more than two or three nuclei. The osteoclastic activity was discontinuous, short periods of vigorous osteoclastic activity being followed by longer periods of inactivity during which the alveolar

crest may be covered with osteoblastic cells or in many cases devoid of bone cells. The net effect in the mono-infected rats is rapid bone loss. Periods of osteoclastic activity were associated with dense inflammatory infiltration of the interdental tissue, ulceration of col epithelium, and the presence of large adherent plaques of *Actinomyces* (Garant 1976a, b).

Furthermore, the amount of bone loss varied from one part of the jaw to the other. Heijl et al. (1980) reported that the most pronounced bone loss occurred in the interproximal areas, especially in the maxillary jaw. Thus, the area between the first and second maxillary molars always showed the most pronounced loss of alveolar bone. The buccal surfaces of the teeth in comparison to the lingual surfaces consistently displayed only minor signs of bone degradation. Listgarten (1978) observed that, although all molar interdental spaces were theoretically equally at risk for the development of periodontal disease, the earliest clinically detectable lesion in the experimental animals was observed in the maxilla, 3 weeks after infection.

The impaction of hair and bedding material was related to periodontal tissue destruction in germ-free rats of the Sprague–Dawley strain and rats mono-infected with microorganisms frequently associated with human gingivitis and periodontitis lesions (Heijl et al. 1980). In fact, only in areas with severe impaction of foreign material could a destruction of the supporting apparatus is recognized. The degree of the impaction was assessed in histological sections according to the following criteria: score 0: no or minor impaction of hair and/or bed bedding material into the interdental area; no evident penetration into underlying epithelium or sub-epithelial connective tissue; score +: large amounts of impacted material present in the interdental area resulting in distortion of soft tissue in an apical direction; epithelial barrier intact, but impacted material has penetrated into the epithelium; and score ++: severe impaction; impacted material present within the sub-epithelial and supracrestal connective tissue. Epithelium often ulcerated (Heijl et al. 1980). In aim to minimize trauma to periodontal tissue from impac-

tion of hair and bedding, Klausen (1991) recommended that the rats are to be kept on large granular size bedding, which does not enter in the gingival sulcus, and the amount of free hair in the environment can be reduced by frequent change of bedding.

It has been showed that diet and bedding conditions have the potential of seriously influencing outcomes of studies of periodontal disease in rats. Regular diet caused significantly more horizontal bone loss ($P=0.0001$) and significantly less periodontal bone support ($P<0.0001$) than the same kind of diet with a smaller grain size. Wood chip bedding in the rats' cages significantly reduced periodontal bone support ($P<0.0001$) compared to a wire mesh floor, and a simultaneous use of regular diet and bedding decreased it even further ($P=0.0023$). Finally, by using finely milled diet, a wire mesh floor, and tap water, instead of conventional breeding methods of regular diet, bedding, and acidic water, it was possible to breed rats with minimal signs of periodontal disease.

Periodontal inflammation has been reported to affect men and women differently, the Sprague–Dawley female rats having a greater risk for inflammatory-based systemic diseases than males (Bain et al. 2009).

Calculus formation can be studied in different strains of rats where diet seems to be more consistent factor. Rats fed a sucrose-rich diet developed a rapid proliferation and overgrowth of bacteria plaque, mainly Gram positive, covering the molar fissures, the interdental spaces and the marginal gingiva. This ultimately resulted in rounded or crater-like gingival pockets (Weinberg and Bral 1999).

21.7.3 Assessment of Periodontal Inflammation in Experimental Periodontitis in Rat

Assessment of periodontal inflammation and alveolar bone loss in experimental periodontitis in rat can be done by different methods: semi-quantitatively by clinical examination under a dissecting microscope (Listgarten et al. 1978),

morphometric measurements of the distance cemento-enamel junction to alveolar bone crest on defleshed jaws (Crawford et al. 1978; Klausen et al. 1989a; Lohinai et al. 1998), radiographic methods (Gaegauf-Zollinger et al. 1982; Nowotny and Sanavi 1983; Katz et al. 1996, 1999; Reymond and Cimasoni 1997) biochemical methods (Chang et al. 1988; Klausen 1991), micro-CT (Seto et al. 2007; Liu et al. 2010; Wilensky et al. 2009; Polak et al. 2010; Rogers et al. 2007; Park et al. 2007) and histological studies (Irving et al. 1974, 1975a; Garant 1976a, b).

21.7.3.1 Histological and Histometric Analysis

The histopathology of periodontal disease was studied in gnotobiotic rats mono-infected with *Eikenella corrodens*, a Gram-negative rod isolated from a human periodontitis lesion. In aim to evaluate the **clinical severity score**, Listgarten et al. (1978) examined the interdental spaces in the molar periodontitis rat region with a dissecting microscope, after the animals were killed by decapitation and heads fixed in a mixture of 5% glutaraldehyde and 4% paraformaldehyde, buffered to pH 7.2 with sodium cacodylate. The periodontal status was scored as follows: 0=no disease: teeth firm, in line and in contact with one another, without any impacted hair or bedding in the gingival sulcus; 1=mild disease: teeth firm, in line and in contact with one another, with evidence of some impacted hair or bedding in the sulcus; 2=moderate disease: teeth firm and in line, open contact points, and definite impaction of hair and bedding in the gingival sulcus; 3=severe disease: teeth out of line, possibly loose, with extensive impaction of hair and bedding, open contacts and alveolar ridge deformities. For each time interval, the periodontal status was assessed by a mean severity score, obtained by adding the scores of all available interdental spaces in the molar regions (8 per animal) and dividing by the number of spaces scored.

The histological and histometric analysis of affected periodontal tissues in experimental periodontitis in rat is the standard method used in most studies and is performed with an image

analysis system under light microscopy. Several landmarks can be determined (**Fig. 21.48**) (Nuñez et al. 2010):

- Length of the defect=distance from the CEJ to the notch.
- Length of the epithelium=distance between the gingival margin (GM) and the apical end of the junctional epithelium (JE) (GM–JE).
- Length of the connective tissue adhesion=distance between the most coronal cementum (CC) and the most apical extent of the JE (CC–JE).
- Length of the CEJ to the osseous crest (OC) (CEJ–OC).
- Length between the CEJ to the GM.
- Proportion of mineralized tissue in the furcation area (i.e., a 1,000- μ m zone under the furcation and between the roots). **Figures 21.49** and **21.50** represent the area of interest (César-Neto et al. 2006).
- Area (mm^2) of bone loss in the furcation region (Almeida et al. 2008; Fernandes et al. 2010; Nociti et al. 2000; Nogueira-Filho et al. 2004).

The histologic method is currently the criterion standard, but it has drawbacks, especially because the pattern of bone resorption in periodontitis is not uniform. In addition, problems during histologic processing, difficulties in carrying it out, its high cost and long-time requirements are well-known disadvantages of histologic analyses (Fernandes et al. 2007).

21.7.3.2 Morphometric Measurements

Various morphometric methods on defleshed jaws were used in aim to assess alveolar bone loss: **the distance method** is widely used (Klausen et al. 1989a; Crawford et al. 1978; Ryder 1980; Evans et al. 1992; Lohinai et al. 1998; Baker et al. 2000b; Arai et al. 2005; Vardar-Sengul et al. 2008). In it, bone loss is taken as the distance from the cemento-enamel junction to the alveolar bone crest as measured at exactly defined sites. **The area method** uses the exposed root surface areas as a measure of bone loss (Burckhardt et al. 1981). The area defined by the cemento-enamel junction and the alveolar bone crest and five vertical distances were measured with a planimeter on prints enlarged from standardized radiographs of defleshed maxillae and mandibles. The verti-

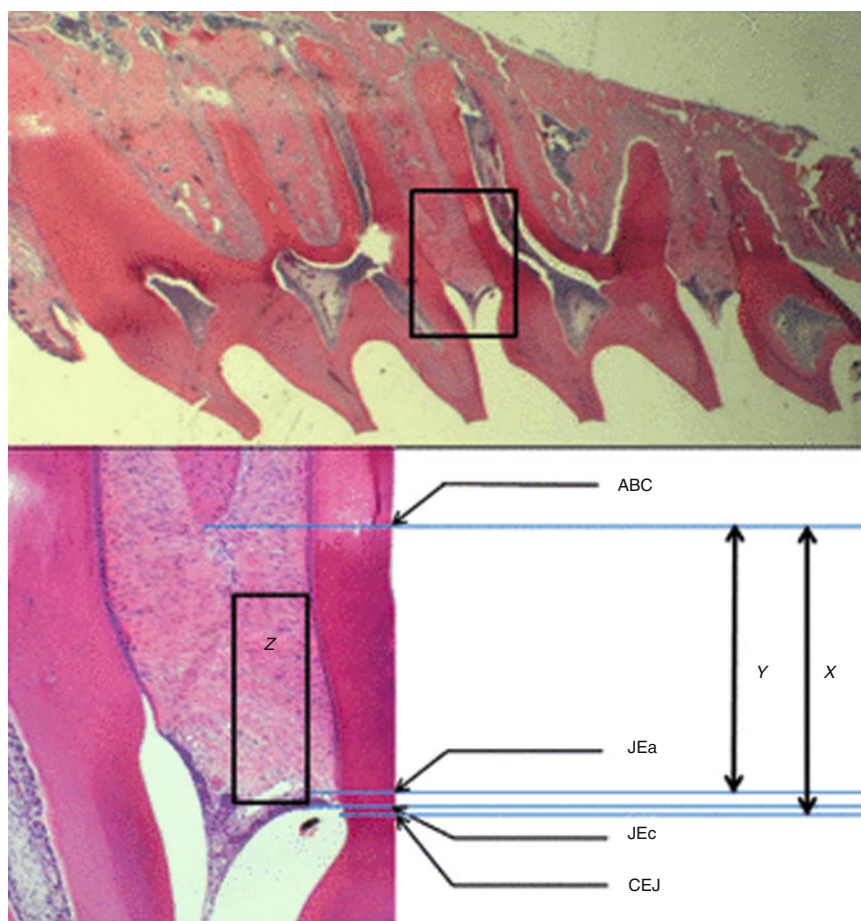


Fig. 21.48 Periodontal histometric measurements performed at the mesial surface of the maxillary second molar in rats. The *top micrograph* shows the histology of the molar region from a control rat. A higher magnification view of the boxed region in that micrograph is shown below. Histometric measurements, including the attachment loss (CEJ to JEc), the alveolar crest bone level (CEJ to $ABC=X$), the connective tissue attachment (JEa to

$ABC=Y$), and the area of inflammatory cell-infiltrated connective tissue ($ICT=Z$) in a zone of 0.14 mm^2 of sub-epithelial gingiva (the boxed zone in the lower micrograph), were performed. *CEJ* cemento-enamel junction, *JEc* the coronal level of epithelial cells, *JEa* the apical level of epithelial cells, *ABC* the coronal level of alveolar bone crest (Cheng et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

cal distances between the cemento-enamel junction and the alveolar bone crest or deepest point of bone resorption were measured at the distal furcation of the first molars and the furcation of the second and third molars. Interdentally, between the first and second and the second and third molars, the distance was measured perpendicularly from a line drawn between the cemento-enamel junction of the two adjacent molars to the deepest point of the interdental pocket (**Fig. 21.51**) (Kuhr et al. 2004).

In a ligature-periodontitis model (silk ligatures tied around the neck of the second maxillary molar teeth), the two methods of assessment of bone loss on defleshed jaws were compared: the distance and the area method. Both methods confirmed that even without the influence of ligatures, the cemento-enamel junction to alveolar bone crest distance increases with age. The area method demonstrates this even more clearly than the distance method. This was probably attributed to the fact that the

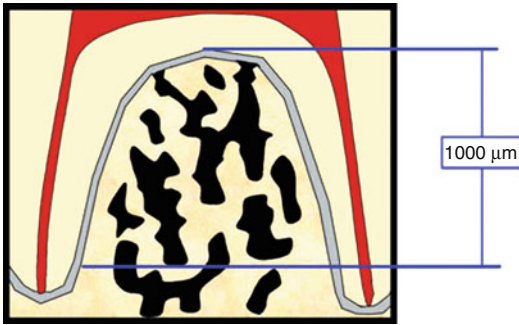


Fig. 21.49 Schematic illustration of the analyzed area (a 1,000 μm-zone under the furcation limited by the roots) (César-Neto et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

area method includes the entire exposed root surface and hence detects changes in bone level much earlier than does the distance method, which only measures certain distances on the root surface. The distance method, however, by virtue of the greater number of distances measured per tooth, allowed a much more exact division of regions into those located within and those outside of the ligature effect zone (Kuhr et al. 2004).

When the two methods were compared, histologic evaluation, measuring bone loss linearly, and morphometric evaluation, measuring bone loss linearly in defleshed maxillae, it was showed

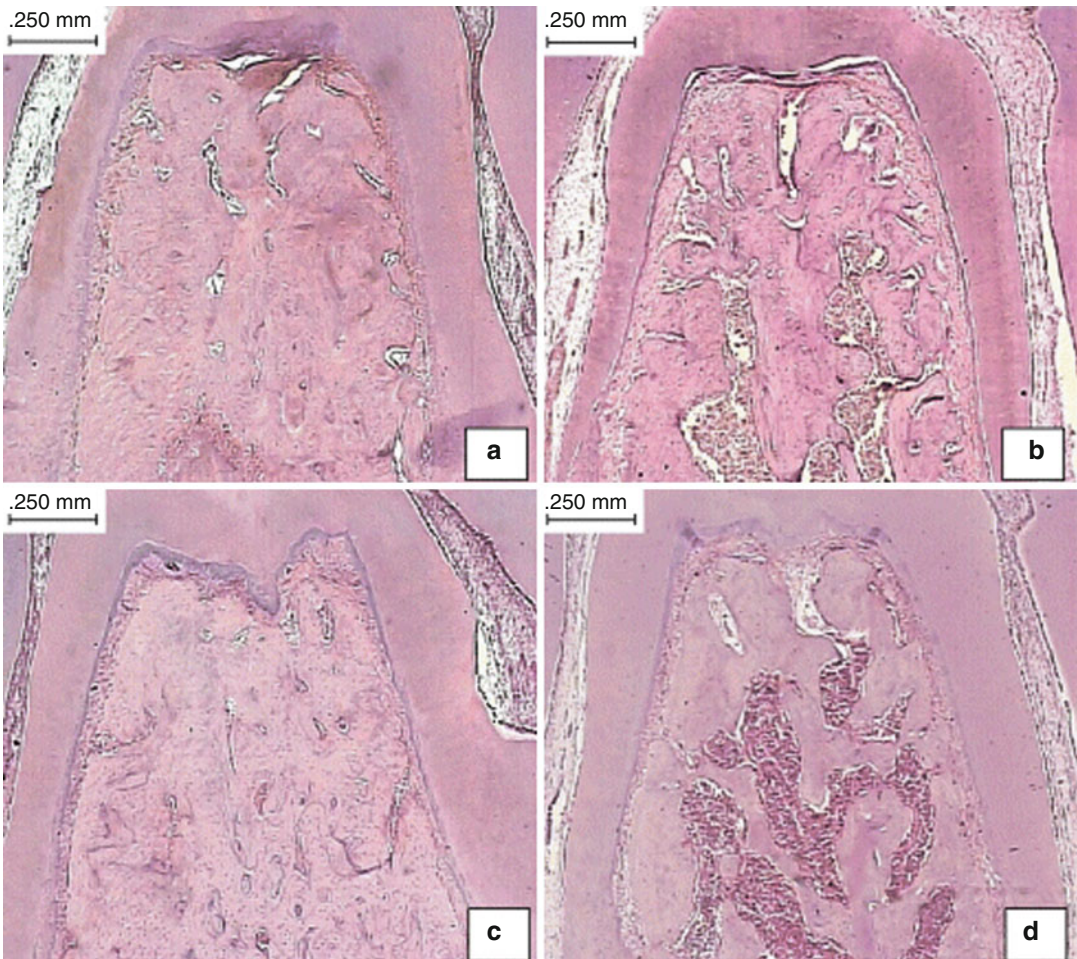


Fig. 21.50 Photomicrographs (a–d) illustrate the histological aspect of the furcation bone for groups 1 (control), 2 (2 months of cigarette smoke inhalation), 3 (3 months of cigarette smoke inhalation and 2 months without expo-

sure) and 4 (5 months of cigarette smoke inhalation), respectively. Original magnification 5x; hematoxylin and eosin (César-Neto et al. 2006. Reprinted with permission from John Wiley & Sons, Inc.)

Fig. 21.51 (a) Distance method. Overview of maxillary posterior dentition. *Right:* oral with sketched-in distances measured. 30×. (b) Area method. Overview of maxillary posterior dentition. *Right:* oral with sketched-in areas measured. 30× (Kuhr et al. 2004. Reprinted with permission from John Wiley & Sons, Inc.)

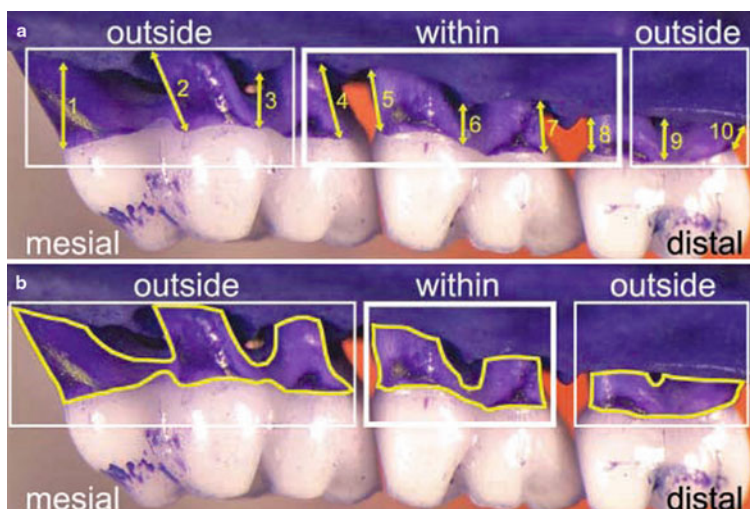
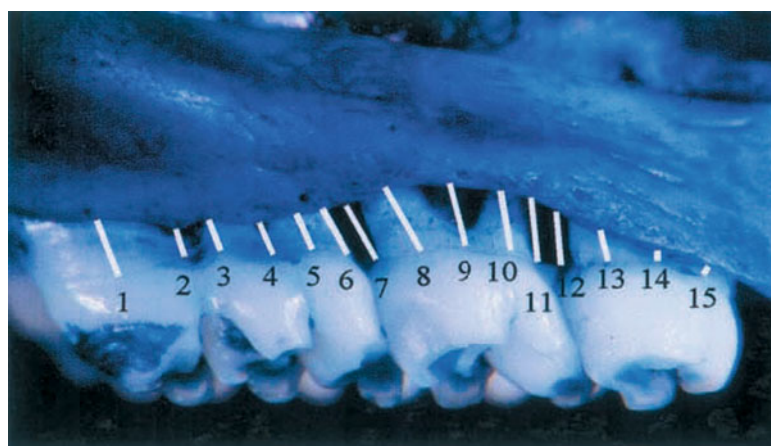


Fig. 21.52 Measurement of periodontal bone loss (PBL). The distance between buccal bone crest and cemento-enamel junction is measured at six sites on the first molar, five sites on the second molar and four sites on the third molar. PBL = sum of the 15 measurements (Björnsson et al. 2003. Reprinted with permission from John Wiley & Sons, Inc.)



that both methods are capable of detecting vertical periodontal bone loss. The comparison between methods, however, revealed statistically significant differences between the groups without ligature when the buccal surface was analyzed and might have been caused by an unwanted inclination of the histologic sections. Despite the care taken, such inclination may have generated an occlusal image of the alveolar bone crest (Fernandes et al. 2007).

21.7.3.3 Radiographic Measurements

The vertical component of bone loss can be evaluated on radiographs (Figs. 21.52 and 21.53) (Gaegauf-Zollinger et al. 1982; Nowotny and Sanavi 1983; Katz et al. 1996, 1999; Holzhausen et al. 2002, 2004; Björnsson et al. 2004):

1. By measuring the area between alveolar bone crest and each CEJ of individual molars.
2. By the mean vertical distance between CEJ and alveolar bone crest. This distance was calculated by dividing the areas of individual molars between CEJ and alveolar bone crest by the mesio-distal tooth width measured either at the level of CEJ or half between CEJ and alveolar bone crest. The resulting values in mm for the corresponding left and right molars in mandibles for maxilla were added.
3. By measuring the vertical distances between CEJ and the deepest point of bone resorption at five defined locations. This point could be located at the level of alveolar bone crest; more often, however, translucent zones were

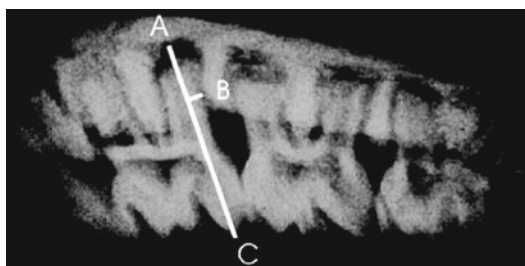


Fig. 21.53 Measurement of periodontal bone support (PBS). The apex (A) of a distal root and the distal cusp tip (C) are located and the distance between A and C (AC) traced and measured in pixels. A line is traced from the deepest bony defect, intersecting AC at a right angle. Finally, the intersection of the two lines (B) is located and the distance from the apex to the intersection (AB) measured in pixels. PBS is calculated according to the formula: $PBS = AB/AC \times 100\%$ (Björnsson et al. 2003. Reprinted with permission from John Wiley & Sons, Inc.)

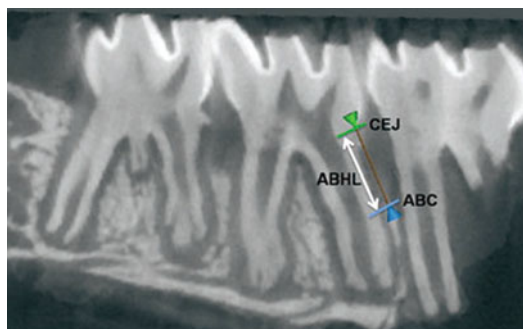


Fig. 21.54 Linear measurement of alveolar bone height loss (ABHL). The ABHL was assessed from images of two-dimensional (2D) micro-computerized tomography (micro-CT) sections. The distances (in mm) from the cemento-enamel junction (CEJ) to the alveolar bone crest (ABC) along the distal and medial root of the second molar were measured (Liu et al. 2010b. Reprinted with permission from John Wiley & Sons, Inc.)

indicative of an intrabony pocket. These distances were measured at the distal furcation of the first molar and the furcation of the second and third molar perpendicularly. Interdentally, between the first and second and third molar, the distance was measured perpendicularly to a line drawn between the CEJ and of the two adjacent molars and the deepest point of the interdental bony crater (Gaegauf-Zollinger et al. 1982). The radiographic results were expressed as the amount of vertical bone loss in millimetres (Katz et al. 1999).

According to Klausen et al. (1989a), the radiographic method to diagnose and measure bone loss is comparable to the morphometric method with respect to reproducibility and ability to discriminate between various experimental groups. However, since the morphometric method mainly measures horizontal bone loss, whereas the radiographic method detects intrabony interproximal defects, it was concluded that future studies would benefit from applying both methods to assess alveolar bone loss in rats.

21.7.3.4 Micro-computerized Tomography (Micro-CT) Measurements

However, the above-mentioned methods provide only linear or two-dimensional (2D) information because they measure horizontal bone loss; they

do not yield data about the three-dimensional (3D) intrabony changes that occur during periodontal infection. Recent studies have shown that **micro-computerized tomography (micro-CT)** is more sensitive in measuring bone mass and bone microstructure than conventional methods (Seto et al. 2007; Liu et al. 2010; Wilensky et al. 2009; Polak et al. 2010; Rogers et al. 2007; Park et al. 2007). The images reconstructed by micro-CT scanning provide 3D information about internal hard tissues. Thus, by measuring bone height, bone density, bone structure and other parameters using micro-CT, detailed information about bone changes over time can be obtained (Liu et al. 2010).

Micro-CT can elaborate cross-sectional tomograms of $\approx 10 \mu\text{m}$ thick and then allows:

- *Linear (2D) measurements* of alveolar bone height loss (ABHL) taken (in mm) from the cemento-enamel junction (CEJ) to the alveolar bone crest along the distal and medial root of the second molars (**Fig. 21.54**) (Seto et al. 2007).
- *Volumetric micro-CT measurements* can be carried out following the selection of a 3-D ROI. Although artificial landmarks could have been placed prior to the scanning of the images, it was more convenient and more reproducible to use given morphologi-

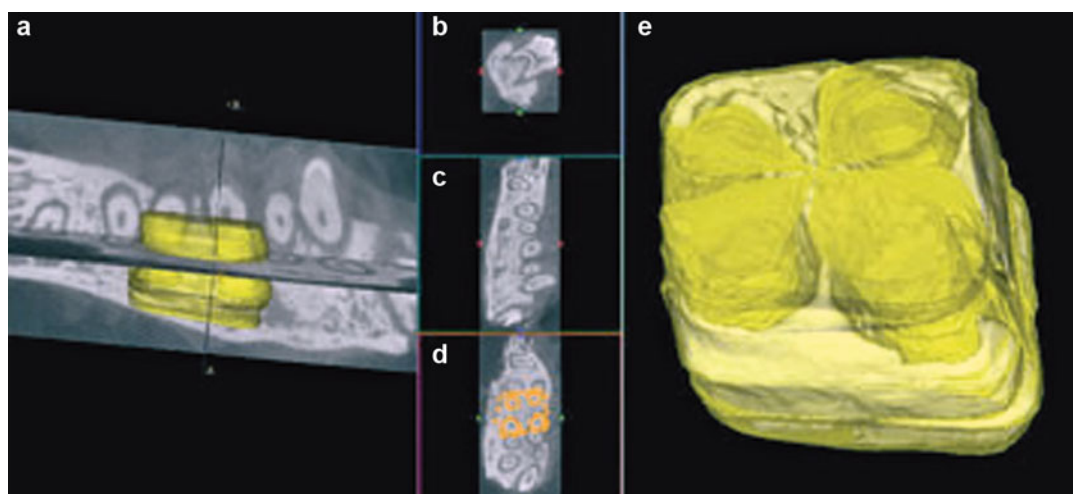


Fig. 21.55 Method for creating three-dimensional (3D) regions of interest (ROI) used in volumetric measurement of alveolar bone loss. (a) 3D ROI tools were selected. Border lines were drawn along the roots of the second molar using landmarks of the most distal root of M1 (d-M1) and the most medial root of M3 (m-M3) from the roof of the furcation to the root apex ((b) proximal–distal;

(c) buccal–lingual; (d) medial–lateral), then a 3D ROI was generated by the software based on the resultant two-dimensional (2D) contours. (e) Alveolar bone to be analyzed was constructed following the border of the 3D ROI (Liu et al. 2010b). Reprinted with permission from John Wiley & Sons, Inc.)

cal features (e.g. outer root prominences and root furcations) (Park et al. 2007). Next, a 3D ROI was generated using the micro-view image analysis software based on the resultant 2D contours (Figs. 21.55 and 21.56). Finally, parameters in the volume of selected and contour-drawn ROIs were estimated: bone mineral density, bone volume fraction, bone trabecular thickness, trabecular separation, trabecular number (Seto et al. 2007).

21.7.3.5 Biochemical Measurements

The major disadvantage of the above-mentioned techniques is that they only measure the total tissue destruction that has occurred until the day of sacrifice; they do not give information about disease activity (ongoing tissue destruction at the time of sampling). For that purpose, biochemical methods are valuable, such as the activity of gingival collagenase and cathepsins (Klausen 1991), myeloperoxidase activity, an indicator of polymorphonuclear leucocyte (PMN) accumulation

(Di Paola et al. 2007) and matrix metalloproteinases expression in experimental periodontitis (Vardar-Sengul et al. 2008; Ramamurthy et al. 2002b; Achong et al. 2003).

21.7.4 Advantages and Disadvantages for Using Rats in Periodontal Research

Rats are cost-effective, easy to handle and allow for the standardization of experimental conditions in genetically similar individuals (summarized in Table 21.7). Rats are suitable for study of the effects of systemic diseases and pharmacological therapies on tissue destruction and regeneration and for evaluation of physiological alterations related to ageing. In addition, tissue destruction and regeneration in an immunodeficient background can be investigated in this model. Surgery and the evaluation of study endpoints are challenging, however, because of the animals' small size. Moreover, the rodent

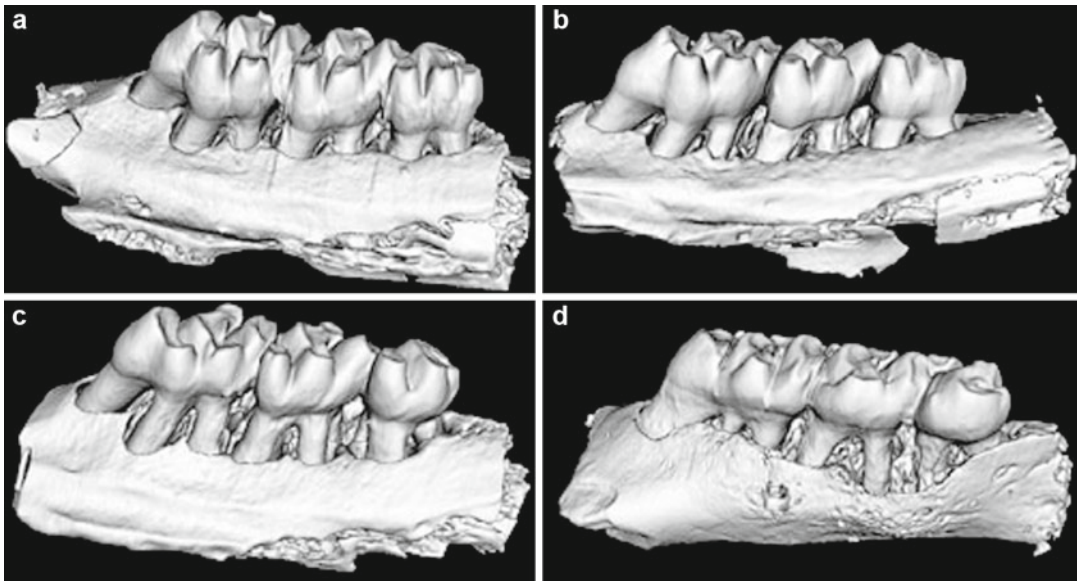


Fig. 21.56 Surface-reconstructed images of the alveolar bone on day 28. Representative images are shown from (a) the unligated side of saline group A, (b) the ligated side of group A, (c) the ligated side of group B (daily administration of 0.83 mg/kg of nicotine) and (d) the ligated side of group C (daily administration of 1.67 mg/kg of nicotine) (Liu et al. 2010b. Reprinted with permission from John Wiley & Sons, Inc.)

Table 21.7 Advantages and disadvantages of pre-clinical rodent models for oral reconstruction Pellegini et al. (2009). Reprinted with permission from Sage Publications

Defect type	Advantages	Disadvantages
Fenestration (periodontal)	• Gives a proof-of-concept in a short time frame • Well-contained defects • No gingival tissue ingrowth	• Narrow healing time window • Small size, surgical microscopes required; technically challenging • Rapid repair as kinetic healing model • Cannot measure healing at junctional epithelial-connective tissue interface • Not a ‘natural disease’ model with microbial influence
Capsule (vertical ridge)	• Standardized shape-dimension • Well-contained defect from the local environment • Contained regenerative response with cells and tissue emanating from the residual bone	• Not applicable to alveolar bone • Isolation from oral environment • Ectopic bone healing model (bone repair beyond normal bony envelope)
Alveolar socket (tooth extraction)	• Easy and fast to perform • Well reproduces the events occurring in bone healing	• Small size • Rapid bone repair compared to human • Non-critical size (kinetic) defect • Mandibular extractions technically demanding
Infrabony peri-implant defect	Surgically created • Short time-frame needed for defect generation • Standard morphological dimension	• Rapid bone repair • Narrow evaluation time window (kinetic defect) • Surgical defect without microbial influence

dentition undergoes continuous tooth eruption, including bone and cementum apposition, which has to be considered when a study is planned (Pellegrini et al. 2009 and references therein).

21.8 Hamster

The dentition formula of hamsters is identical to rodents: I 1/1, C 0/0, Pm 0/0 and M 3/3. As in rats, the molars move with time following the growth of the jaws and occlusal wear. The shift in the occluso-distal direction and the continuous eruption of molars seem to be less evident than in rats (Struillou et al. 2010).

The type of periodontal disease hamsters develop is similar to rats in that there is primarily gingival retraction with horizontal bone loss, the interdental septum being too narrow to induce infrabony defects. Inflammation is not a prominent feature, as it is in humans (Weinberg and Bral 1999).

Experimental periodontitis development is very similar to rats and can be generated in hamsters in different ways:

- Using silk ligature 4 weeks accompanied by the application of *P. gingivalis* (Hojo et al. 2008)
- Administration for 12 weeks of hyperglucidic Keyes 2,000 diet (UAR, France) which promotes bacterial plaque growth and subsequent periodontitis (Baron and Saffar 1978, 1990, 1991; Escartin et al. 2003; Yamazaki 1992)
- With indigenous microorganisms such as *Actinomyces* (Jordan and Keyes 1964; Jordan et al. 1972; Irving et al. 1974)

In addition, hamsters have buccal pouchers (HBP) lined with stratified squamous epithelium that are useful for studying oral carcinoma (Oz and Puleo 2011). Multiple signalling pathways are dysfunctional in both human and hamster oral squamous cell carcinomas. In particular, cell proliferation, apoptosis and angiogenesis are intricately interlinked in malignant transformation of the HBP mucosa. The HBP carcinogenesis model is the best-known animal system for intervention by chemopreventive agents because of easy accessibility for examination and follow-up of lesions. A number of synthetic and natural

products have been documented to exhibit chemopreventive efficacy in the HBP model (Nagini 2009).

21.9 Mice

Mice have been extensively used to study aspects of periodontal disease, including the inflammatory host response and induction of alveolar bone loss (Liang et al. 2010; Graves et al. 2008). In this regard, mice constitute a relatively inexpensive and convenient model; there is a wealth of information on their immune system, and, importantly, the availability of lines of transgenic mice with targeted gene deletions offers important mechanistic tools (Liang et al. 2010; Graves et al. 2008). These tools can be used to dissect proximal and downstream events in periodontal pathogenesis, such as mechanisms of microbial sensing and activation of specific proinflammatory pathways. In contrast, causal mechanistic relationships between suspected etiological factors and periodontal breakdown cannot normally be addressed in human studies owing to important ethical considerations (Liang et al. 2010; Graves et al. 2008).

The dental formula of mice is typical rodent dentition: I 1/1, C 0/0, Pm 0/0 and M 3/3. Incisors have continuous growth, and molars present complex physiological modifications with ageing. There is occlusal wear, bucco-occlusal motion and high hypercementosis at the apical part of each root (Struillou et al. 2010).

Several mouse strains have been used in periodontal studies, although BALB/c mice are considered the model of choice owing to their increased susceptibility to infection-induced periodontal bone loss (Fig. 21.57). Since sham-infected mice in these models do not typically develop appreciable periodontal bone loss, this might give the impression that mice do not develop naturally occurring periodontitis, that is, induced by their own oral flora. In fact, the reason that sham-infected mice do not develop obvious signs of periodontitis is probably due to their young age (usually used when 8–12-weeks old), consistent with the fact that periodontitis is normally associated with advanced age mice that

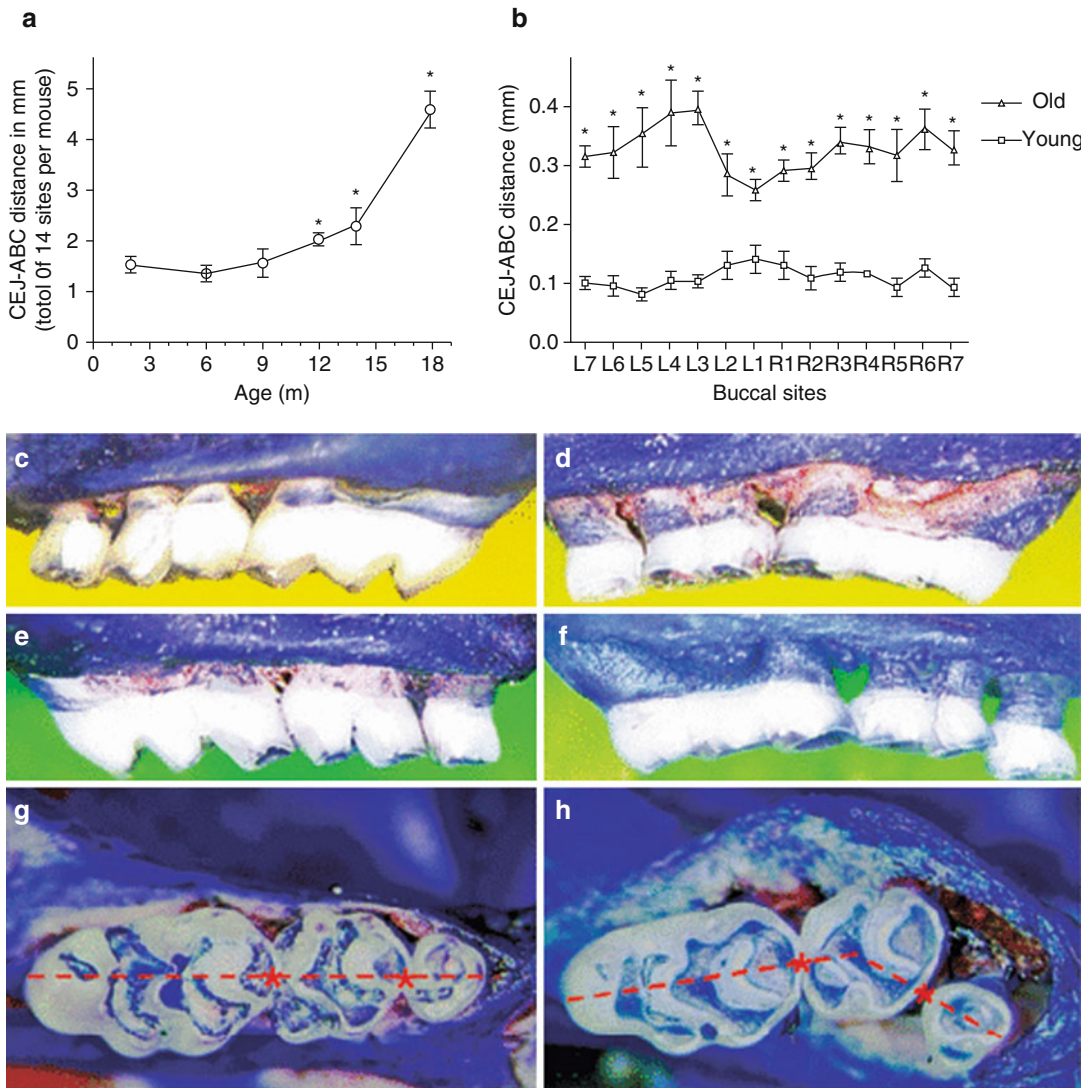


Fig. 21.57 Periodontal bone loss as a function of age in BALB/c mice. (a) Young mice (8–10 weeks of age), old mice (≥ 18 months of age) and mice of intermediate ages (6, 9, 12 and 14 months old) were used to determine their periodontal bone levels. The distance (in mm) from the cemento enamel junction (CEJ) to the alveolar bone crest (ABC) was measured at 14 predetermined maxillary buccal sites, and the readings were totaled for each mouse. The data are means $\pm$ SD ($n=5$ mice). (b) Analytical data from the two extreme age groups (young vs. old). Each point corresponds to a measured site (L1–L7, left maxilla; R1–R7, right maxilla) and represents means $\pm$ SD ($n=5$ mice). Asterisks denote significant differences ($P<0.05$) in CEJ–ABC distances compared with young mice (the greater the CEJ–ABC distance, the greater the bone loss).

The experiment was repeated, with additional sets of 5 mice per group yielding consistent results. (c–f) Representative images from the maxillae of young (c right; e left) and old mice (d right; f left). Extensive areas of resorbed alveolar bone are evident in the old mice. (g, h) Maxillary molar blocks seen from the occlusal surfaces of young (g) and old mice (h). Note that all three molars in young mice are in line with each other (i.e. a straight line can connect their contact points, indicated by asterisks). Owing to tooth mobility, this relationship did not apply to the molar blocks of old mice, many of which displayed overt migration of molars (especially of the second molar in the buccal direction) (Liang et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

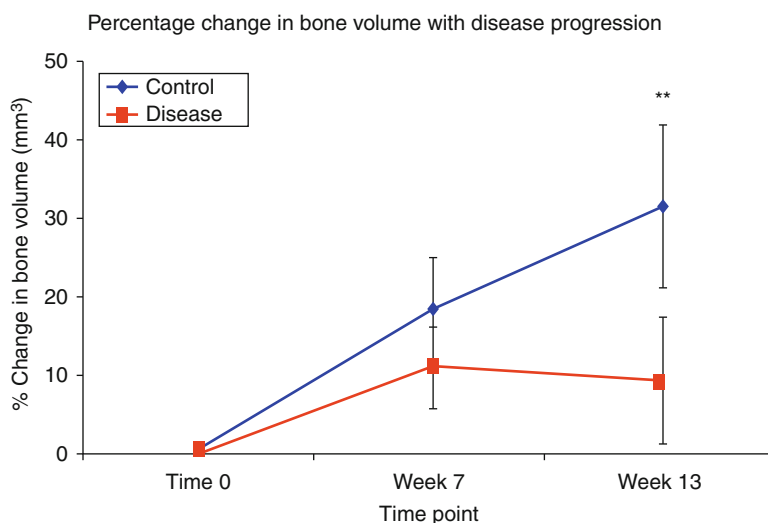


Fig. 21.58 Average percentage change in bone volume (mm^3) for diseased mice (*Porphyromonas gingivalis* inoculated; $n=5$) compared with controls (no disease or treatment; $n=5$) for the following three time-points: time 0, weeks 7 and 13. Percentage changes are shown from time

0 to week 7 and from time 0 to week 13. Bars represent mean $\pm$ SEM. The Student's *t*-test was used to determine significance. ** $P<0.01$, statistically significant compared with the control at week 13 (Cantley et al. 2009). Reprinted with permission from John Wiley & Sons, Inc.)

develop naturally induced periodontal bone loss as a function of age. In comparison with young mice, old mice displayed significantly ($P<0.05$) increased periodontal bone loss, accompanied by elevated expression of proinflammatory cytokines (interleukin-1 β , tumour necrosis factor- α and interleukin-17 α) and innate immune receptors involved in the induction or amplification of inflammation (Toll-like receptor 2, CD14, CD11b, CD18, complement C5a receptor and triggering receptor expressed on myeloid cells 3) (Liang et al. 2010).

Several models of experimental periodontitis were used in mice.

21.9.1 Oral Gavage Model of Experimental Mouse Periodontitis

The introduction of human strains of bacteria by oral gavage and subsequent impact on the periodontium has used several periodontal pathogens such as *P. gingivalis* (Baker et al. 1994, 1999, 2000a, b; Lalla et al. 1998; Alayan et al. 2006; Cantley et al. 2009; Wilensky et al. 2009; Wong et al. 2010; Takano et al. 2010; Sasaki et al. 2010;

Barros et al. 2011; Hernández et al. 2011), *A. actinomycetemcomitans* (Garlet et al. 2006, 2010; Repeke et al. 2010), *Tannerella forsythia* (Sharma et al. 2005) and *Porphyromonas gulae*, an animal periodontal pathogen considered to be equivalent to human *P. gingivalis* (Hardham et al. 2005), *F. nucleatum* alone or associated with *P. gingivalis* (Polak et al. 2009; Bendyk et al. 2009).

Typically, specific pathogen-free BALB/cByJ male mice were first treated with sulfamethoxazole-trimethoprim (10 ml/l water) for 10 days ad libitum (to reduce the oral flora) followed by a 5-day antibiotic-free period. Mice are infected by gavage with 10^9 cfu/ml of live bacteria in 100 μ l of PBS with 2% carboxymethyl cellulose three times at 48-h intervals (Baker et al. 1994). In many reports, mice are euthanized 6 weeks after the final infection dose. Substantial bone loss typically takes more than 4 weeks to develop and can be measured histologically, by morphometric analysis or by micro-computed tomography. For bone loss assessment, the investigators most often examine the bone around the maxillary molars because induction of bone loss in the mandible is slower, whereas incisors are not included in the assessment due to their continuous eruption (Figs. 21.58, 21.59 and 21.60)

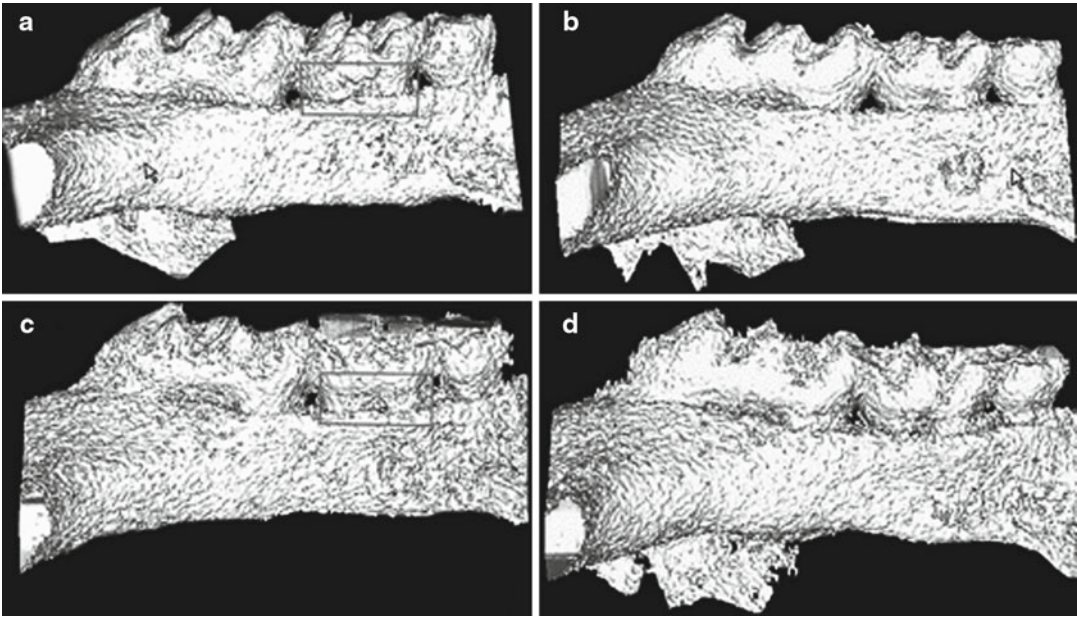


Fig. 21.59 Three-dimensional models (8.7 voxels) of the three molars of the maxilla (created using the Skyscan analysis program ANT) demonstrating the bone resorption in diseased mice and drug-treated mice compared with the control. The box in *panels a and b* shows the same area around molar no. 2 in which bone resorption became evident as the disease progressed. (a) Three-

dimensional model of a diseased mouse at time 0. (b) Three-dimensional model of a diseased mouse at week 13. (c) Three-dimensional model of a week 13 control mouse. (d) Three-dimensional model of a week 13 drug-treated mouse (Cantley et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

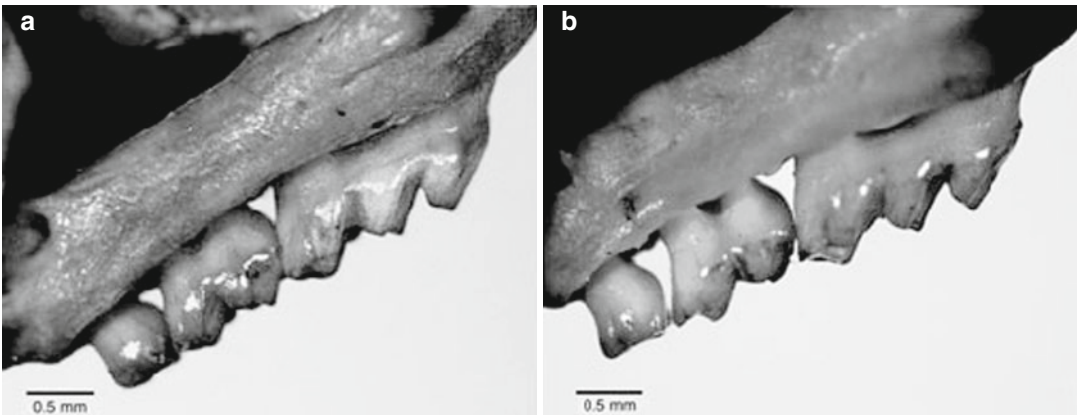


Fig. 21.60 Stereo images of mouse heads showing the three molars for one diseased mouse (*Porphyromonas gingivalis* inoculated) compared with a control mouse (no disease or treatment). Images were taken using a stereomicroscope (MZ16FA) with a $\times 1.0$ objective lens. Scale:

280 pixels/mm. (a) Control mouse right-hand side lingual view. (b) Diseased mouse (*P. gingivalis* inoculated) right-hand side lingual view (Cantley et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

(Graves et al. 2008). The oral gavage model is useful for addressing a wide variety of hypotheses related to periodontal pathogenesis, ranging from the role of the host response to

virulence traits of suspected periodontopathogens and the interconnection of those factors with systemic parameters (Table 21.8) (Graves et al. 2008).

Table 21.8 Hypotheses tested in the oral gavage model (Graves et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Genetic basis for host susceptibility or resistance to periodontal disease	
Different genetic mouse strains display differential susceptibility to <i>Porphyromonas gingivalis</i> -induced periodontal bone loss	Baker (2005)
<i>Il1b</i> , <i>Tnf</i> and <i>Stat6</i> are associated with susceptibility, whereas <i>Il15</i> and <i>Selp</i> with resistance to <i>P. gingivalis</i> -induced periodontal bone loss	Hart et al. (2004)
Role of specific cytokines or antimicrobial molecules in periodontal disease	
Increased susceptibility of interleukin 10 (IL-10) knockout mice to <i>P. gingivalis</i> -induced periodontal bone loss	Sasaki et al. (2000)
IL-17-receptor-knockout mice display increased susceptibility to <i>P. gingivalis</i> -induced periodontal bone loss	Yu et al. (2007)
Inducible nitric oxide synthase (iNOS)-knockout mice exhibit increased periodontal bone loss in response to <i>P. gingivalis</i> infection	Alayan et al. (2006)
Identification of potential virulence factors	
Wild-type <i>P. gingivalis</i> , but not a FimA-deficient mutant, induces periodontal bone loss and accelerates atherosclerosis	Gibson et al. (2004)
Diminished induction of periodontal bone loss by <i>P. gingivalis</i> mutant expressing FimA fimbriae lacking the FimCDE accessory components	Wang et al. (2007a)
Diminished induction of periodontal bone loss by <i>P. gingivalis</i> mutants lacking expression of Kgp and/or RgpB	Pathirana et al. (2007)
Diminished induction of periodontal bone loss by <i>T. forsythia</i> mutant lacking BspA adhesin	Sharma et al. (2005)
Immune evasion through exploitation of innate receptors TLR2 deficiency attenuates <i>P. gingivalis</i> -induced periodontal bone loss	Burns et al. (2006)
Blockade of complement receptor-3 inhibits <i>P. gingivalis</i> -induced periodontal bone loss	Hajishengallis et al. (2007)
Periodontal disease connection with systemic diseases	
Oral infection with <i>P. gingivalis</i> accelerates atherosclerosis	Gibson et al. (2004) Lalla et al. (2003)
<i>P. gingivalis</i> -induced periodontal bone loss is enhanced in diabetes	Lalla et al. (1998)
Blockade of RAGE (receptor for advanced glycation end products) suppresses periodontal bone loss in diabetic mice	Lalla et al. (2000)
Proof-of-principle immunization studies to identify candidate vaccine antigens	
Systemic immunization using RgpA and Kgp gingipain peptides protects against <i>P. gingivalis</i> -induced periodontal bone loss	O'Brien-Simpson et al. (2005)
Systemic immunization using intact RgpA protects against <i>P. gingivalis</i> -induced periodontal bone loss	Gibson and Genco (2001)
Oral immunization with FimA-expressing <i>Streptococcus gordonii</i> vector protects against <i>P. gingivalis</i> -induced periodontal bone loss	Sharma et al. (2001)
Potential and future uses of this model	
To determine the role of pattern-recognition receptors in protection or susceptibility to periodontitis using genetically modified mice	
To determine the impact of ageing in periodontal disease susceptibility using young and aged mice	
To investigate the protective potential of specific antagonists of disease-promoting mechanisms (e.g., destructive inflammation, bacterial evasion of immunity)	
To investigate the role of specific bacterial genes in interbacterial co-adherence and biofilm formation	

21.9.2 Calvarial Models

The calvarial or scalp model was developed by Brendan Boyce to study the effect of cytokines on osteoclastogenesis (Boyce et al. 1989). In this model, a stimulus is injected directly into

the connective tissue overlying the calvarial bone in a small volume (e.g. 25 µl) of carrier. Classic inflammatory events occur including the rapid expression of proinflammatory cytokines within a few hours and the recruitment of PMNs within 24 h. Bone resorption is induced within

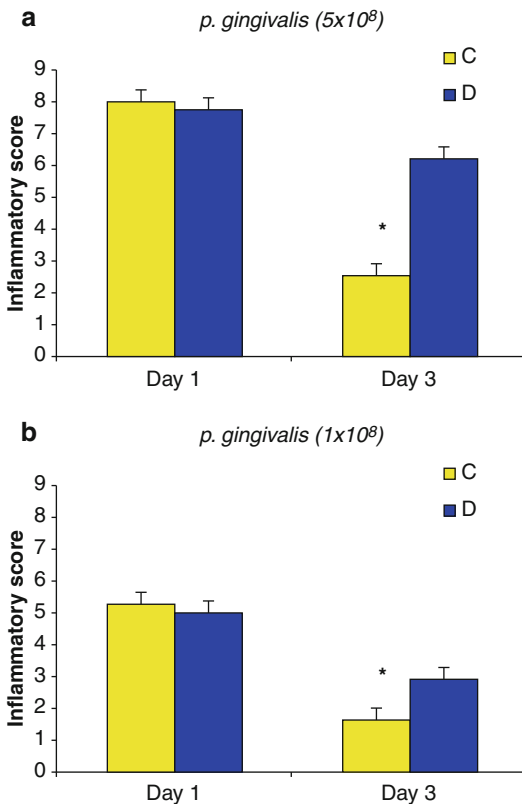


Fig. 21.61 *P. gingivalis* stimulates formation of a more persistent inflammatory infiltrate in diabetic mice. Diabetes was induced by multiple low-dose streptozotocin treatment, while control mice were treated with citrate buffer alone. (a) *P. gingivalis* (5×10^8) or (b) *P. gingivalis* (1×10^8) was inoculated s.c. into the scalp. The degree of PMN infiltration was measured in H&E-stained sections at the centre of the infiltrate. A scale was used ranging from no inflammation to severe inflammation with necrosis (0–9). Results are shown as mean $\pm$ SEM of 6 mice per group. *A significant difference ($P < 0.05$) between the control and diabetic groups (Graves et al. 2005b. Reprinted with permission from Sage Publications)

3–5 days depending on the size of the stimulus (Fig. 21.61). If the stimulus is small, there is an inflammatory event that resolves with relatively little tissue destruction. On the other hand, a large inoculum can create a perturbation large enough to induce a soft tissue wound and considerable bone resorption that subsequently undergoes repair. New bone formation can then be examined as part of the healing process and can be measured 8–12 days following inoculation of the stimulus. Thus, a potentially impor-

tant use of this model is the ability to examine coupling, the process of reparative bone formation that occurs after bone resorption. The use of an outbred mouse strain, CD-1, avoids some of the extremes of inbred strains. Thus, the calvarial model is useful in examining the sequence of inflammation, destruction and repair. This may be particularly important in periodontal disease where the net loss of tissue may be related to loss of balance between destruction of bone and connective tissue and the degree of repair that occurs (Graves et al. 2008 and references therein). A summary of hypotheses that have been tested in the calvarial model and potential uses of the model are described in Table 21.9.

21.9.3 Murine Back Abscess Model (Airpouch/Chamber Models)

To produce an airpouch, 5 cm³ of air is injected subcutaneously into the back of a mouse or a rat near the midline. A second injection is usually administered after 3 days to maintain the patency of the airpouch. After 6 days, the airpouch will have become filled with fluid and it is lined by epithelium. The model was used extensively in the rheumatoid arthritis (RA) literature to investigate the dynamics of synovial inflammation and the initiation of inflammation. The shortcoming of the airpouch model is the duration of the experiment; it is not practical to conduct a chronic inflammation experiment with the airpouch, with maintenance of the airpouch being the limiting factor. This problem was solved in part by the use of a tissue cage and chamber models. The chamber is a coiled stainless-steel wire that is surgically implanted subcutaneously in the back of the mouse (or other animal), and the wound is allowed to heal for 10 days. Like the airpouch model, the interior of the coil becomes epithelialized, allowing for the injection of bacteria or other substances into the lumen of the chamber. Fluid can be aspirated from the chamber at intervals, or the entire chamber can be excised and histology can be performed (Graves et al. 2008). The use of the airpouch and chamber mice models has been useful to test specific pathogenic hypotheses of periodontal disease (Genco et al.

Table 21.9 Hypotheses tested in the calvarial model (Graves et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

Impact of the host response on hard and soft tissue damage	
Interleukin 1 (IL-1) induces osteoclastogenesis and bone resorption	Boyce et al. (1989)
Periodontal pathogens induce osteoclastogenesis and bone resorption partly through induction of prostaglandins	Zubery et al. (1998)
<i>Porphyromonas gingivalis</i> induces an inflammatory infiltrate and osteoclastogenesis indirectly via TNF	Graves et al. (2001)
Periodontal pathogens <i>P. gingivalis</i> and <i>Aggregatibacter actinomycetemcomitans</i> induce greater cytokine expression than a commensal bacterium, <i>Streptococcus gordonii</i>	Kesavalu et al. (2002)
Activation of the acquired immune response significantly increases expression of innate immune cytokines and an inflammatory infiltrate stimulated by <i>P. gingivalis</i>	Leone et al. (2006)
Bacterial virulence factors	
Low dose <i>P. gingivalis</i> LPS induces osteoclast formation through IL-11 and prostaglandins	Li et al. (2002a)
High dose <i>P. gingivalis</i> LPS induces osteoclasts through IL-1 and TNF activity	Chiang et al. (1999)
<i>P. gingivalis</i> fimbriae are pro-inflammatory but do not play an essential role in the innate immune response to <i>P. gingivalis</i> in connective tissue	Graves et al. (2005a)
<i>P. gingivalis</i> LPS, fimbriae and intact bacteria exhibit a qualitatively different stimulation of the innate immune response	Zhou et al. (2005)
LPS indirectly stimulates fibroblast apoptosis through TNF activity	Alikhani et al. (2003)
Impact of systemic disease (diabetes)	
Diabetes types 1 and 2 both cause prolonged inflammation in response to <i>P. gingivalis</i>	Graves et al. (2005b), Naguib et al. (2004)
Diabetes enhanced apoptosis limits the repair of a bacteria induced wound	Al-Mashat et al. (2006)
Diabetes enhanced TNF increases fibroblast & osteoblast cell death initiated by <i>P. gingivalis</i> , which significantly interferes with repair of connective tissue and bone	Liu et al. (2006a)
Diabetes interferes with the coupling of bone resorption and formation following bacteria induced bone loss	Al-Mashat et al. (2006)
Potential and future uses of this model	
To investigate the role of specific virulence factors by genetic manipulation of bacterial genes in stimulating or evading a host response in connective tissue <i>in vivo</i>	
To investigate the role of specific genes in mounting an innate immune response to oral bacteria in connective tissue <i>in vivo</i> using genetically modified mice	
To investigate the role of specific genes in mounting an acquired immune response to oral bacteria in connective tissue <i>in vivo</i> using genetically modified mice	
To investigate host mechanisms or systemic factors that modulate repair following bacteria induced injury to connective tissue	
To investigate mechanisms or systemic factors that affect osseous repair following bacteria induced bone resorption	

1991, 1992; Hachicha et al. 1999; Hu et al. 2006; Pouliot et al. 2000; Schenkein 1989; Wilensky et al. 2009).

The strength of the airpouch and chamber models in small animals is clearly the whole-animal aspect for the measurement of host-response parameters. The strength of the model is enhanced, particularly in mice, by the availability of geneti-

cally modified knock-out or transgenic animals that can be used in the experiments. For short-term, acute inflammation experiments, the strength of the airpouch model is its simplicity and ease of use with no requirement for surgical interventions. The major advantage of the chamber models is the ability to perform long-term experiments that can be adapted to immunization

and chronic inflammation and pathogenesis experiments (Graves et al. 2008).

21.9.4 Ligature-Induced Periodontal Model

P. gingivalis-soaked ligature approach provides a simple and direct route to deliver enough bacteria into the murine gingival sulcus to colonize and to initiate pathogenesis of periodontitis. The ligatures were replaced every other day to maintain a constant-specific bacterial infection. Before infection, *P. gingivalis* was not isolated in any of the mice, irrespective of their group. Furthermore, *P. gingivalis* was not recovered in untreated control mice and only recovered in the infected mice (day 7). These data suggested that *P. gingivalis*-soaked ligatures can successfully deliver *P. gingivalis* into mouse gingival sulcus and thereby contribute to the induction of experimental periodontitis: significantly increased epithelial downgrowth ($P<0.05$), inflammation ($P<0.05$), alveolar bone loss ($P<0.05$) and osteoclast activity ($P<0.05$). The alveolar bone loss in a group receiving ligature only was significantly less than that in the *P. gingivalis*-soaked ligature group 10 days after ligature placement (0.240 ± 0.015 vs. 0.403 ± 0.036 mm, respectively; $P=0.008$). Furthermore, noticeable alveolar bone loss induced by ligature-only animals results from gingival trauma with subsequent mixed murine oral plaque colonization—a process significantly different than what is observed in human disease. Finally, this process requires considerable time to develop (Li and Amar 2007).

A similar approach was used by Watanabe et al. (2011). To induce periodontitis, 8-0 silk sutures were placed around maxillary second molars. *Escherichia coli* LPS (2.5 ng in phosphate-buffered saline) was soaked into the mesial and distal interproximal portion of ligatures, and ligature placement was confirmed weekly from weeks 2 to 8. General anaesthesia was given for placing ligatures and for weekly confirmation of ligature placement. Control animals were also

administered general anaesthesia to control for the effects of anaesthesia. To assess the severity of induced periodontitis, the alveolar bone loss as the area (mm^2) bordered by the cemento-enamel junction, the crest of alveolar bone and the mesial and distal line angles on the buccal and lingual sides of maxillary second molars was measured. Ligature placement with LPS treatment for 10 weeks resulted in significant bone loss compared with the control mice (Fig. 21.62) (Watanabe et al. 2011).

21.9.5 Chemically Induced Mouse Model

An alternative method for inducing inflammation of oral tissues is by using trinitrobenzene sulphonc acid (TNBS) or dextran sulphate sodium (DSS) (Oz and Puleo 2011; Oz et al. 2010; Oz and Ebersole 2010). Dextran sodium sulphate acts to undermine the epithelial barrier and as an immune cell activator, resulting in innate immune damage to the tissues. Trinitrobenzene sulphonc acid appears to function as a hapten to modify autologous proteins and induce a T cell-mediated response, resulting in autoimmune-like inflammatory responses. Mice of the BALB/c strain were treated biweekly with dextran sodium sulphate (2% DSS) in the diet followed by 1 week of abstinence, repeated for a period of 7–18 weeks. Trinitrobenzene sulphonc acid (TNBS, 2.5 mg solution) was delivered via a micropipette tip (100 μl orally twice a week. TNBS treatment, which primarily mimics the induction of an autoimmune type of destructive inflammation, and DSS treatment, which undermines the integrity of the epithelium, as well as activating various inflammatory cells, both elicited significant bone loss. Both TNBS and DSS caused a localized action on periodontal tissues, with alveolar bone loss observed in both maxilla and mandibles, with progression in a time-dependent manner. Bone loss was detected as early as week 7, with more severe periodontitis increasing over the 18 weeks (Figs. 21.63 and 21.64) (Oz and Ebersole 2010).

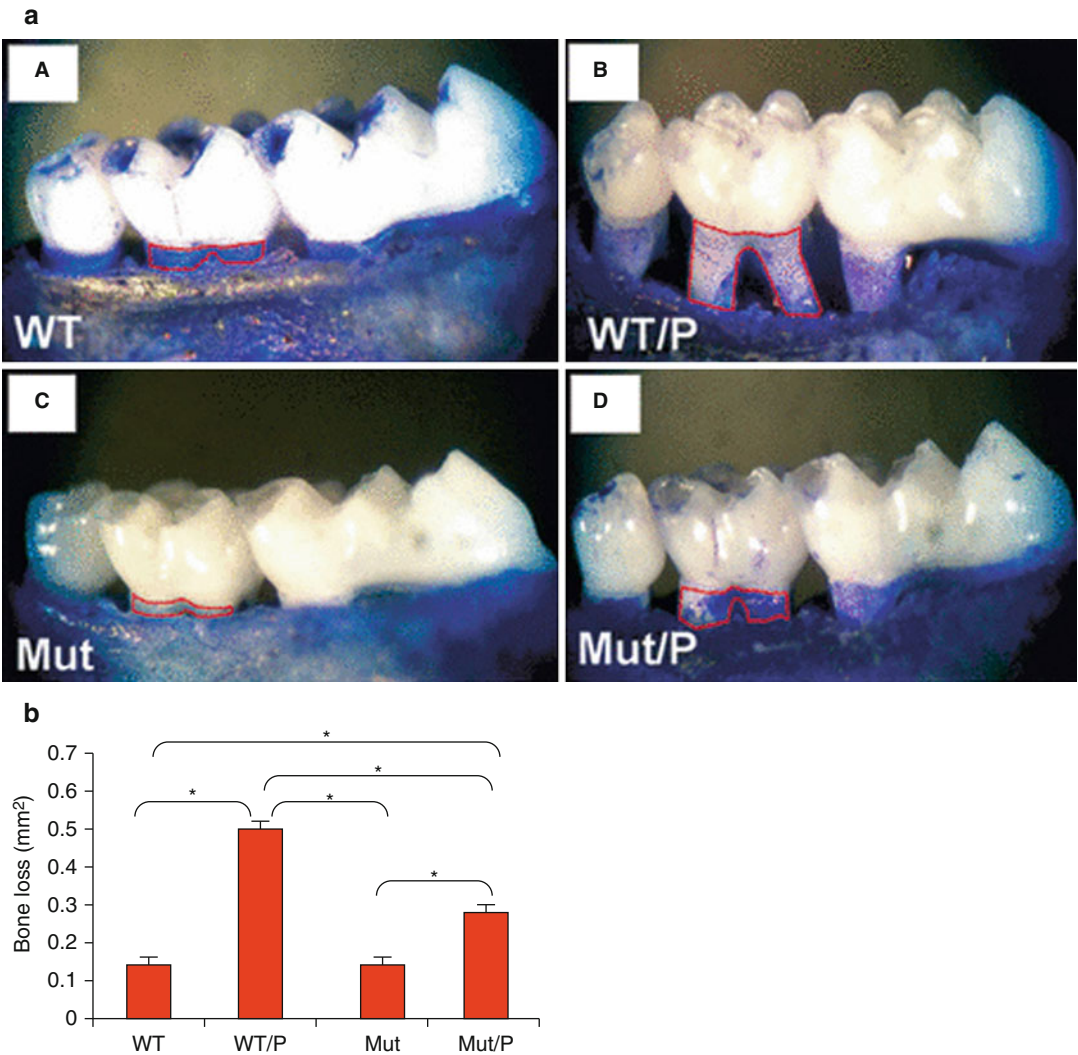


Fig. 21.62 (a) Alveolar bone loss captured by stereomicroscopy of defleshed mouse maxilla. (A) Maxilla from a control wild-type (WT) mouse without periodontitis; (B) maxilla from a WT mouse with periodontitis (WT/P); (C) maxilla from a toll-like receptor 4 (TLR4) mutant mouse without periodontitis (Mut); (D) maxilla from a Mut mouse with periodontitis (Mut/P). Red lines are drawn on the mesial and distal transition line angles. The area within the line was calculated as bone

loss. The areas were measured from the buccal and palatal sides, and the total area of bone loss was calculated per tooth. (b) Average bone loss per tooth calculated using software (ImageJ). Data are presented as mean $\pm$ SEM. Statistical analysis was performed using a Mann–Whitney *U*-test. * $P < 0.05$ between direct comparison groups (Watanabe et al. 2011) (Reprinted with permission from John Wiley & Sons, Inc.)

21.9.6 Gum Pocket Mouse Model

An abscess in a gum pocket, resulting from bacterial infection, is a common source of chronic halitosis. Although antibiotics are generally

prescribed for abscesses, they require multiple treatments with risks of creating resistant bacterial strains. Liu et al. (2010a) had developed a novel vaccine using ultraviolet-inactivated *Fusobacterium nucleatum* (*F. nucleatum*), a

Fig. 21.63 Significant bone resorption at the levels of maxillary, mandibular and total quadrant after 18 weeks of treatment with TNBS or DSS ($P<0.001$) (Oz and Ebersole 2010. Reprinted with permission from John Wiley & Sons, Inc.)

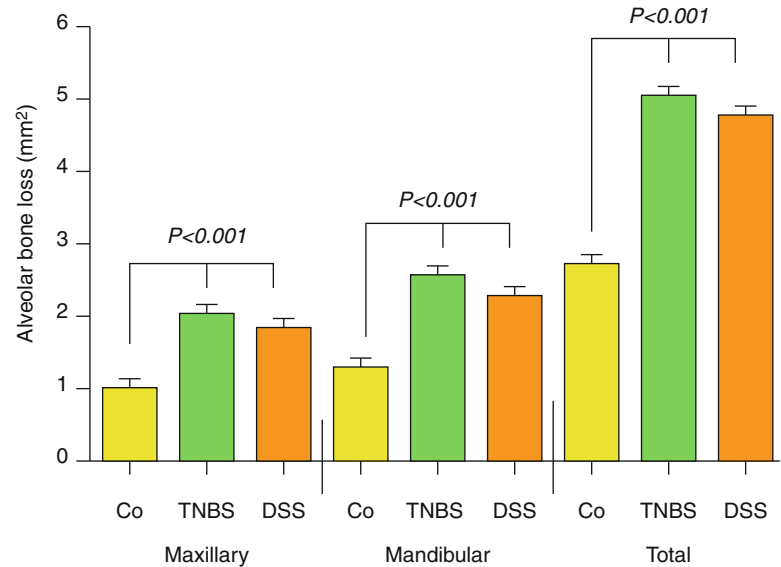
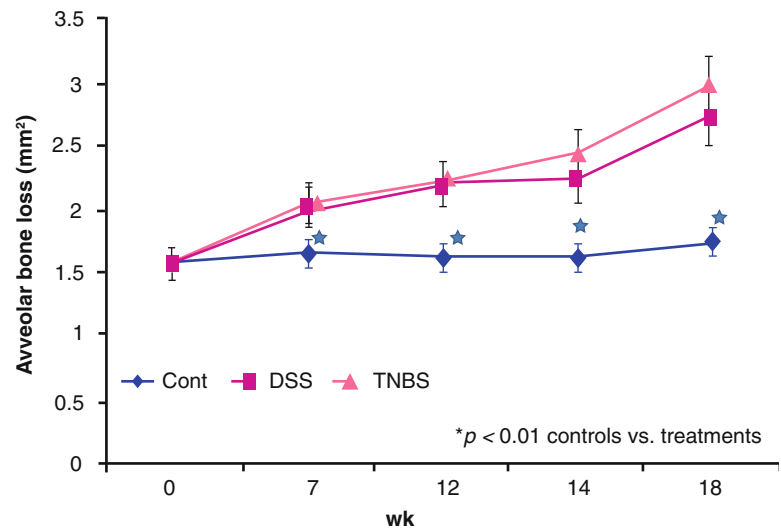


Fig. 21.64 Extent of the alveolar bone resorption. Significant mandibular bone loss was detected as early as week 7 and progressed to a more severe periodontitis by week 18. The y-axis denotes area of bone loss (in mm²) ($n=5$ /group) (Oz and Ebersole 2010. Reprinted with permission from John Wiley & Sons, Inc.)



representative oral bacterium for halitosis. A gum pocket model, established by continuous inoculation of *F. nucleatum*, was employed to validate the vaccine potency (Fig. 21.65). Mice immunized with inactivated *F. nucleatum* effectively minimized the progression of abscesses, measured by swollen tissues of gum pockets. Most notably, the immunized mice were capable of eliciting neutralizing antibodies against the

production of volatile sulphur compounds of *F. nucleatum* (Liu et al. 2009).

21.9.7 Murine Vaccine Studies

Murine models are extensively used in studies assessing humoral immune factors and the effects of immunizations (see Table 21.10) (Persson 2005).

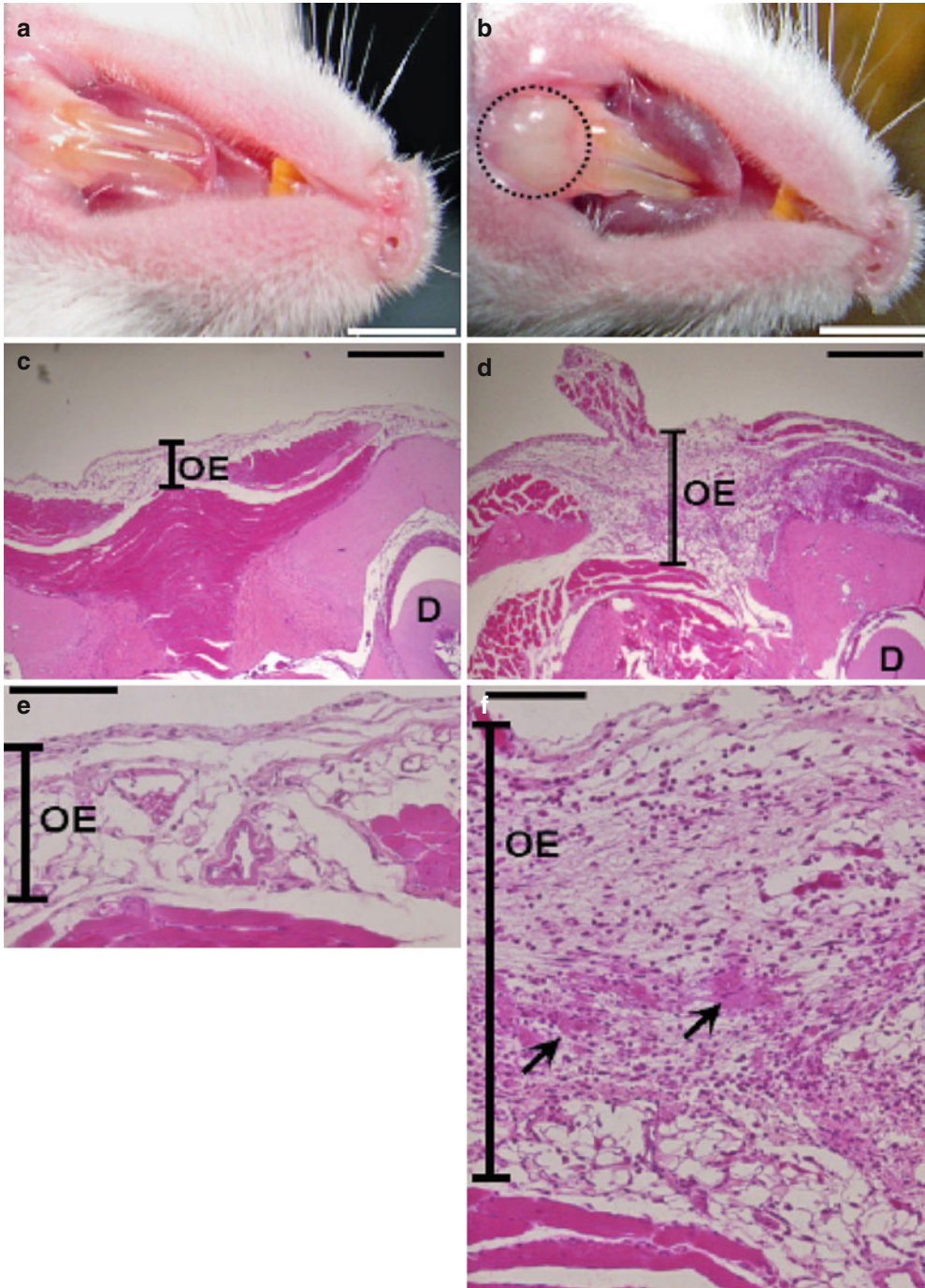


Fig. 21.65 An abscess induced by the injection of live *F. nucleatum* into the gums of lower incisors of ICR mice. One hundred microlitre of PBS (**a, c, e**) or *F. nucleatum* (4×10^8 CFU in PBS) (**b, d, f**) was inoculated into a mouse oral cavity for 3 days, in order to induce abscesses. An aliquot of 30 μ l was injected into the gums of lower incisors with a 28-gauge needle. An aliquot of 30 μ l was directly dropped into the oral cavity. The remaining 40 μ l of ali-

quot was spread over the surface of the tongue. An abscess (**b, circle**) was observed in the gums 3 days after injection with *F. nucleatum*, but not with PBS (**a**). H&E magnification 4 $\times$ (**c, d**) and magnification 20 $\times$ (**e, f**) indicated that granulomatous inflammation (**arrows**) occurs in an abscess. *D* dentin, *OE* oral epithelium. Bars (**a, b**)=2.2 mm. Bars (**c, d**)=1 mm. Bars (**e, f**)=0.5 mm (Liu et al. 2009. Reprinted with permission from Elsevier)

Table 21.10 Examples of murine study models and results from vaccine trials (Persson 2005) (Reprinted with permission from John Wiley & Sons, Inc.)

Study	Antigen	Study type	Results	Conclusion
Gemmell et al. (2004)	<i>F. nucleatum</i> ATCC 25586 <i>P. gingivalis</i> ATCC 33277 Viable bacteria used in immunization schedule	27 BALB/c female mice Various immunization schedules and a placebo group (intra-peritoneal injections once per week (4 weeks))	1. ELISA and chemiluminescence assays 2. Production of both IgG1 and IgG2 subclass antibodies and higher levels for the combination vaccine <i>P. gingivalis</i> / <i>F. nucleatum</i>	A vaccine candidate may be enhanced by combinations of <i>P. gingivalis</i> and <i>F. nucleatum</i>
Ross et al. (2004)	<i>P. gingivalis</i> YH522, ATCC 33277, ATCC53978 12 clinical isolates from Seattle, Norway, Romania, Sudan	Genetics research and animal vaccine study with constructed gene vaccine using knockout mice challenged to <i>P. gingivalis</i> infection	1. The gene and expressions of PGO695 and PGO694 are soluble and potentially useful as vaccine candidates 2. Gene expression is highly conserved across <i>P. gingivalis</i> species 3. Immunization provides reduction in bone lesions	Additional work needed to improve results to enhance solubility. There is a loss in epitope domains when using truncated versions of proteins as vaccines
DeCarlo et al. (2003)	<i>P. gingivalis</i> ATCC 33277 HA2 sequence (Genbank PGU68468.1) cloned with <i>E. coli</i>	Fischer CD F(344) rats Test group recombinant HA2 (eight animals) Placebo group (Freunds' adjuvant) (eight animals) three immunizations	1. Sham-immunized animals developed no anti-rHA2 IgG2 antibodies, whereas rHA2-immunized animals did develop IgG antibodies measurable over 70 days. The extent of protection against bone loss dependent on titre 2. and absence of Freunds' adjuvant in vaccine	Protective effect against bone loss in the absence of adjuvant in the vaccine
Rajapakse et al. (2002)	Whole-cell killed <i>P. gingivalis</i> ATCC 33277/ ATCC53978 + adjuvant Rgp (ArgX and LysX proteinase) + adjuvant	Four groups (two of them (sham immunized) of Sprague-Dawley rats two immunizations with 3-week intervals After antibiotics treatment challenged with <i>P. gingivalis</i> infection	Immunization with the RgpA-Kgp 1. <i>P. gingivalis</i> W50-induced high IgG titres Immunization restricted colonization with 2. <i>P. gingivalis</i> ATCC 33277 Cross-reactivity between ATCC 33277 and W50 (ATCC53978) 3. Immunization prevents bone loss	Immunization with Kgp39 and Rgp44 may prevent or reduce periodontitis in humans
Gibson and Genco (2001)	<i>P. gingivalis</i> A7A1–A28	BALB/c mice Gingipain PgpA and RgpB Whole-cell heat-killed <i>P. gingivalis</i> three immunizations	1. Immunizations with gingipains PrgA and PrgB stimulate <i>P. gingivalis</i> serum IgG antibodies Purified denaturated PgpA and RgpB identified 2. 5 polypeptide bands identified as 45, 44, 17, 15, 27 kDa Immunization with RgpA but not RgpB protects against 3. <i>P. gingivalis</i> -induced bone loss	Whole-cell heat-killed <i>P. gingivalis</i> and gingipain RgpA (haemagglutinin domain) provide protection against bone loss

Table 21.10 (continued)

Study	Antigen	Study type	Results	Conclusion
O'Brien-Simpson et al. (2000)	<i>P. gingivalis</i> ATCC 33277 <i>P. gingivalis</i> W50	BALB/c mice immunized with Rgp-Kgp proteinase adhesion complex of ATCC 33277 or W50 with or without adjuvant (IFA) Abdominal injections	1. When challenged with W50 post-immunization, those immunized against W50 showed significantly smaller abdominal lesions When challenged 2. with ATCC 33277, no lesions were found post-immunization Protection 3. against <i>P. gingivalis</i> infection may be mediated via Fc receptor-dependent phagocytosis	RgpA-Kgp protein-adhesin complex protects against <i>P. gingivalis</i> challenge in the murine model
Katz et al. (1999)	<i>P. gingivalis</i> ATCC 33277, 381, A7A1-A28 Hagb gene from 381 was cloned in a pET vector and expressed in <i>E. coli</i>	Fischer CDF 344 rats were immunized with recombinant Hagb plus Freund's adjuvant subcutaneously and then orally exposed to fresh <i>P. gingivalis</i> at days 13 and 14 post-immunization	1. No elevation in salivary IgA as a result of immunization Serum IgG 2. elevated in immunized + infected and in infected-only animals Supernatants from rHagb immunized stimulated lymphoid cell cultures 3. with high levels of interferon, followed by IL2, IL-1, and then IL-4 consistent with T-helper type 1 (Th1) and Th2 responses rHagb reacted with <i>P. gingivalis</i> ATCC 33277, 381, A7A1-A28, and W50 with a 50 kDa protein band representing rHagb	Immunization with purified rHag B induces protective immunity against <i>P. gingivalis</i> infection and provides a potential vaccine candidate against chronic periodontitis in humans
Genco et al. (1998)	<i>P. gingivalis</i> A7436 and HG66-extracted cysteine protease (Gingipain R)	148 BALB/c female mice: non-immunized Peptide A Peptide D Whole-cell Gingipain R1 (95 kDa) Gingipain R2 (50 kDa)	1. When challenged with 1. <i>P. gingivalis</i> , non-immunized and peptide A immunized develop ulcerated lesions, weight loss 2. Gingipain R immunized were protected from abscess formation 3. Reduction in viable. <i>P. gingivalis</i> counts 4. Antisera from. <i>P. gingivalis</i> HG66 recognize gingipains from many different <i>P. gingivalis</i> strains	Immunization with gingipain R provides protection against <i>P. gingivalis</i> infection in mice
Evans et al. (1992)	<i>P. gingivalis</i> 381, 2561 and <i>P. gingivalis</i> ATCC 33277 fimbriae	Germ-free Sprague-Dawley rats (6 groups of 8 rats per group) Sham-immunized non-infected, Sham-immunized infected Whole-cell heat-killed <i>P. gingivalis</i> Purified 43 kDa protein Purified 75 kDa protein Combined 43 and 75 kDa vaccine	43 kDa immunized were protected from alveolar bone loss 75 kDa 2. immunized had no protection against bone loss The combination 43, 3. and 75 kDa provided protection Gingival fluid collagenase activity 4. reduced in 43 kDa immunized animals	43 kDa fimbrial protein may be useful as a vaccine candidate against periodontitis

21.10 Shrew (*Suncus murinus*)

Suncus murinus, a rat-size laboratory house musk shrew, has received attention as a valuable animal model due to ease of handling. Dentition pattern is I3/1, C1/1, P2/1, M3/3 (Takata et al. 1999; Yamanaka et al. 2007). Variation in the number of teeth, including both congenital and postnatal absence, was observed in wild populations of *Suncus murinus* (Jogahara et al. 2008).

S. murinus in captivity shows a high rate of periodontal disease, such as spontaneous gingival swelling with accumulation of plaque, was observed in more than two-thirds of animals older than 200 days. Morphometric analysis of alveolar bone demonstrated a pattern of bone loss that correlated closely with animal age. Histologically, periodontal lesions varying from gingivitis to periodontitis, similar to those observed in humans, were noted. Marked infiltration of lymphocytes and plasma cells in the connective tissue was noted, usually not seen in periodontal lesions of rodents. Although osteoclastic alveolar bone resorption was noted, active bone resorption was not a frequent feature in specimens obtained from chronic inflammatory lesions. Ultrastructurally, degradation of collagen fibres in the inflamed area and ingestion of collagen fibrils by fibroblasts in the deeper connective tissue were often seen (Takata et al. 1999).

Takata et al. (1999) indicated the potential utility of *Suncus murinus* as a model to study periodontal disease, for example, chronic nature of the inflammatory periodontal lesions, similar to those in humans, as well as other advantages including size and ease of handling and housing of these animals.

21.11 Rabbit

The rabbit model of periodontitis was shown to be a relevant model in which the physiology and pathology of periodontal tissues resemble humans with respect to proinflammatory and anti-inflammatory mechanisms (Serhan et al. 2003).

Experimental periodontitis was induced in rabbits with different methods:

- Silk sutures tied around the mandibular second premolars bilaterally, followed by the topical application of 10^9 colony-forming units (CFU) of *P. gingivalis* over a 6-week period (Serhan et al. 2003; Jain et al. 2003; Hasturk et al. 2006a, b, 2007a, b, 2009; Hasturk et al. 2006b). As opposed to other animal models of periodontitis, ligature placement alone does not cause any change in alveolar bone (Hasturk et al. 2012). Local *P. gingivalis* administration in addition to ligature placement led to significant bone resorption and increased inflammation compared with ligature alone (Hasturk et al. 2007a). The histological sections showed disrupted connective tissue layers with irregular fibre arrangement. Numerous blood vessels and inflammatory cells were localized adjacent to the basal layer in the connective tissue. Dense inflammatory infiltration spreads to the lamina dura of the alveolar bone, leading to bone destruction, and the alveolar borders were extremely ragged. Conversely, the ligature-alone group did not show any measurable soft and hard tissue destruction with the absence of *P. gingivalis*, indicating that bacterial challenge is necessary to initiate the host response (Hasturk et al. 2009).
- Teasing cotton pellets into the labial pockets of the lower incisor of rabbits three times a week for 7–24 weeks (Rizzo and Mitchell 1966). The technique was modified by Mallison et al. (1988) who tied the cotton gauze into the gingival sulcus by a horizontal mattress suture of 4-0 silk tied in the sulcus at the base. This secured the cotton square, provided a matrix for bacterial growth, and induced chronic inflammation in the gingiva. The cotton square remained in the sulcus for 15–20 days. An inflammatory infiltrate with numerous plasma cells, characteristic of chronic inflammation, was observed as early as 15 days after implantation of the cotton (Mallison et al. 1988).
- The surgical fenestration defect (a 4×6-mm bone window) in rabbits is less suitable for studies of regeneration of periodontal ligament as a new experimental model to study periodontal regeneration. Up to 6 weeks, the fenes-

tration defects healed partly by repair and partly by regeneration. After 6 weeks, the root had erupted to such an extent that the original root defect shifted into the oral cavity. This signifies that the periodontal ligament bordering the original bone defect site was newly formed during the natural eruption process and not locally regenerated. At 12 weeks, no signs of surgery were present anymore (Oortgiesen et al. 2010).

The periodontal lesion in rabbits' evaluation can be performed with morphometric analysis of the bone loss, histologically or radiologically (Jain et al. 2003). It was showed that the etiologic agents of rabbit periodontal infections consist of a mixture of anaerobic Gram-negative rods, especially *F. nucleatum*; anaerobic Gram-positive nonspore-forming rods, predominantly *Actinomyces* spp.; and aerobic Gram-positive cocci, particularly the *Streptococcus milleri* group (Tyrrell et al. 2002).

The traditionally most commonly used models in rabbits for testing biomaterials are the transcortical drilled holes creating tibial or radial critical-size femoral defects (Aaboe et al. 1994; Schmitt et al. 1997; Le Guehennec et al. 2005; Hautamäki et al. 2008; Monjo et al. 2010; Calvo-Guirado et al. 2011). These defects in long bone are far from the specific situation of periodontal diseases but appear as a very interesting model for testing the bone healing (Struillou et al. 2010).

21.12 Sheep Model

The dentition of the sheep's mandible consists of three incisors (I1–3), a very small canine (C), three premolars (PM1–3), and three molars (M1–3). Sheep incisors have similar size and shape as their human equivalents. Lingual to incisors and canines, a so-called lower dental pad, is present, which consists of bulky, fibrous tissue (Figs. 21.66, 21.67 and 21.68) (Al-Qareer et al. 2004).

'Broken mouth' is the term used to describe premature tooth loss in sheep (i.e. periodontal disease of sheep). The problem of broken mouth is extremely common, with about 70% of all ewes

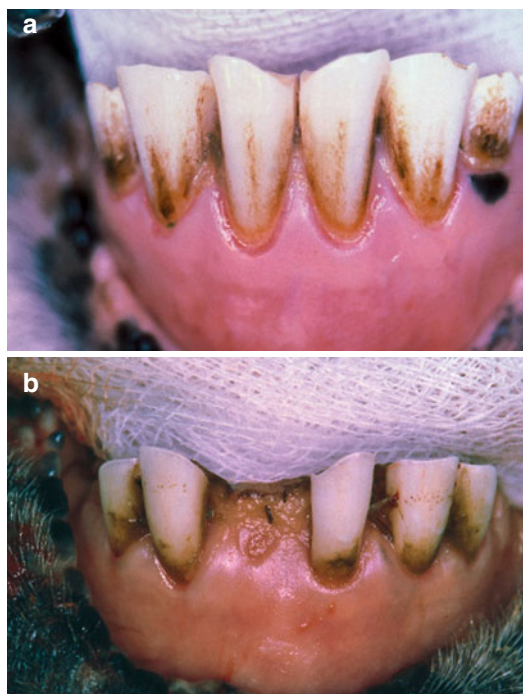


Fig. 21.66 Clinical presentation of sheep anterior teeth. (a) A healthy (HS) animal and (b) a periodontitis-affected (PDS) animal (Duncan et al. 2003. Reprinted with permission from John Wiley & Sons, Inc.)

between 3 and 8 years of age presenting evidence of the condition at slaughter. In sheep, on farms that are free from broken mouth, gingivitis persists throughout the life of the animal. In animals that go on to develop broken mouth, gingivitis worsens, followed by gingival recession, loss of attachment and alveolar bone resorption. The precise aetiology of the progression to destructive periodontal disease remains poorly understood (Kirkham et al. 1991).

The oral microbial flora in periodontally diseased sheep from many countries is consistent with findings from humans with periodontitis, *P. gingivalis*, *B. forsythus*, *F. nucleatum* and *P. intermedia* being consistently present in sheep (Friskien et al. 1989; Ismaiel et al. 1989; McCourtie et al. 1990; Dreyer and Basson 1992; Duncan et al. 2003). In sheep with periodontitis, bone resorption, collagen breakdown, degradation of blood capillary vessel walls, cellular infiltration of plasma cells, lymphocytes and epithelial

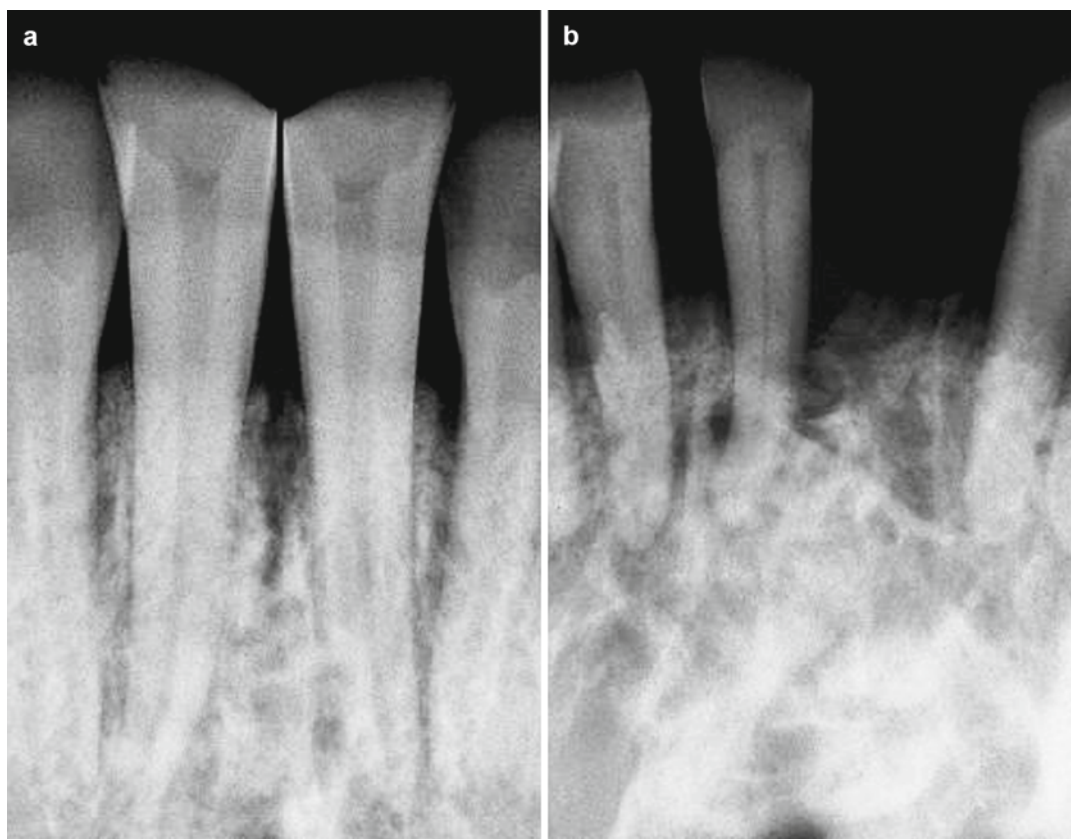


Fig. 21.67 (a, b) Intra-oral radiographs of sheep anterior teeth. (a) A healthy sheep (HS) showing horizontal alveolar bone pattern. (b) A diseased sheep (PDS) showing

horizontal bone loss, angular defects and tooth loss (Duncan et al. 2003. Reprinted with permission from John Wiley & Sons, Inc.)

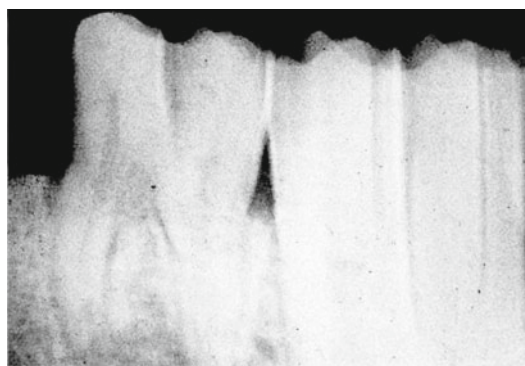


Fig. 21.68 Periodontitis affecting posterior teeth in sheep. Note vertical defect at the distal of the molar and mesial horizontal bone loss (Duncan et al. 2003. Reprinted with permission from John Wiley & Sons, Inc.)

infiltration were significantly greater than in sheep without periodontitis, and the features are similar to those in rapidly destructive forms of periodontal disease in man. The levels of serum IgG antibody reactive against *B. gingivalis* antigens (measured by ELISA) were significantly higher in sheep with periodontitis than those without, similar to the association reported in some types of human periodontal diseases (Ismail et al. 1989).

The sheep has a mandibular second premolar tooth with a bifurcation of similar proportions of that of a small human mandibular molar, thus allowing the application of materials and techniques similar to those presently used in human GTR therapy (Danesh-Meyer et al. 1997).

21.13 Ferrets

The ferret is a carnivore belonging to the Mustelidae family, which also includes the skunk, the otter, and the mink. There are two ferret species, the North American black-footed ferret (*Mustela nigripes*) and the domestic ferret (*Mustela putorius furo*). Two varieties of the domestic ferret are recognized, the fitch ferret, which has a brown coat and a mask-like face, and the albino or English ferret, which has yellowish-white fur and pink eyes (Fisher et al. 1993).

The ferret has both a deciduous and permanent dentition. The permanent dentition consists of incisors; canine; second, third and fourth premolar; and first and second molar (Weinberg and Bral 1999).

The normal histologic features of the periodontal tissues were observed in the domestic ferret. The periodontal ligament width, the cementum layer at the most apical portion of the root, and the distance between the cemento-enamel junction and the alveolar bone crest of the maxillary teeth presented normally greater mean values than these values of the mandibular teeth. The dimension between the cemento-enamel junction to the alveolar crest varied depending on the group of teeth analyzed. The largest distance was 1.3 mm in the mesial aspect of the mandibular canine. The premolars and molars ranged 0.30–0.62 and 0.29–0.53 mm, respectively (Fisher et al. 1993).

The composition of the normal oral mucosa and gingival sulcus microflora of the domestic ferret was to a large extent composed of facultative anaerobic Gram-positive rods and Gram-negative rods, with a preponderance of *Pasteurella* sp. and *Corynebacterium* sp. The environment of the normal dentogingival area does not seem to enhance heavy colonization of anaerobic bacteria, since these groups could not be observed in any sites investigated. As a whole, the composition of the microflora of the ferret presented few similarities to that described in dogs and monkeys. However, the anaerobic portion of the gingival sulcus and oral mucosa microflora related to ligature-induced periodontitis presented also microorganisms, *Porphyromonas* and *Fusobacterium*, observed in experimental periodontitis both in

dogs and monkeys as well as in human periodontal diseases (Fischer et al. 1994).

Ferrets were found to have spontaneous periodontal disease as well as experimentally induced periodontitis (Weinberg and Bral 1999). Experimental ligature-induced periodontal breakdown was provoked in ferrets (Fischer and Klinge 1994). On days 14 and 28, pockets depth and periodontal attachment loss in the experimental teeth were significantly increased as compared to the corresponding values at the control teeth. Histologically after 7 days of ligature in place, a slight bone resorption was observed, but no loss of connective tissue attachment was seen. After 14 days of the presence of the silk ligatures, both bone resorption and connective tissue attachment loss were observed. Root resorption was observed at 14, 21 and 28 days, but more marked on day 28. Fibrocytes/fibroblasts decreased and PMNs increased in numbers in the experimental teeth, as compared with the corresponding control mean values (Fischer and Klinge 1994).

Ferrets are a suitable model for studying calculus formation and inhibition. Harper et al. (1990) indicated that the hard dental deposits removed from the ferrets structurally resembled human calculus, with a typical ‘coral reef’ surface. X-ray diffraction analysis indicated that the mineralized deposits are composed primarily of hydroxyapatite (Harper et al. 1990). Mann et al. (1990) had evaluated the reduction of calculus accumulation in domestic ferrets with two dentifrices containing pyrophosphate, indicating that the ferret model may have substantial utility in the development of additional calculus control products or procedures, as well as in basic studies on calculus formation.

The ferret offers several advantages over other animal models. In contrast to rodents, sequential oral examination can be carried out in individual animals, facilitating studies on the kinetics of calculus formation. Furthermore, ferrets may be used in multiple sequential studies and may be concurrently scored for other oral parameters, such as gingivitis and plaque accumulation, which is not possible in rodents (Harper et al. 1990). However, they can easily escape from standard cages, and they need special maintenance (Oz and Puleo 2011).

21.14 Cat

The morphologic findings in the naturally occurring gingivitis and periodontitis of domestic cats were analyzed by Reichart et al. (1984) and Gavin (1970). In comparison with experimentally induced periodontitis in other animals, periodontal disease involving external root resorption seemed to occur spontaneously in the cat. Clinical inspection of the teeth before histology preparation revealed plaque and calculus accumulation on facial surfaces of 32% of premolars and molars; it was only seen occasionally on the other teeth. Probing at the cemento-enamel junction revealed defects in 25% of all premolars and molars; cavities were located on facial surfaces in most instances. Radiographic signs of alveolar bone destruction were seen in 77.3% of all premolars and molars. Horizontal and vertical bone resorption with destruction of the inter-radicular bone was observed (Reichart et al. 1984).

The bacteriological analysis of subgingival plaque samples obtained from cats showing different stages of periodontal disease revealed that the tendency with higher gingival index scores, and with the most-affected sites, was towards a microbial population composed to a greater extent of anaerobic Gram-negative rods. The most common organism was an anaerobic Gram-negative rod in the black-pigmented *Bacteroides* group that was biochemically similar to *B. gingivalis* but had catalase activity. The black-pigmented *Bacteroides* group and *Peptostreptococcus anaerobius* were found in increasing numbers with increasingly severe periodontal disease. *Pasteurella multocida* was isolated from most samples and appeared to decrease in numbers with increasing periodontal disease (Mallonee et al. 1988).

Domestic cats are unusual in that resorption of permanent teeth by osteo(onto) clasts is a very common problem: epidemiological reports suggest that up to 72.5% of animals suffer from the condition and the incidence of disease increases with age. Resorption is prevalent without evidence of clinical disease and occurred at younger ages than previously reported. It can initiate anywhere on the root surface, but lack of repair of lesions at the cemento-enamel junction indicates

that mechanisms of replacement are absent or compromised in this region. Whereas resorption of the root may undergo repair, resorption at the cervix may progress to clinically evident lesions. The reason for the buccal side bias of feline osteoclastic resorptive lesions is not well understood, but it has been suggested that feline osteoclastic resorptive lesions may be a sequel of periodontal disease, and a buccal side bias in periodontal inflammation and plaque accumulation has been described in feline periodontitis. Self-cleansing of lingual tooth surfaces by tongue movement is the most likely explanation (DeLaurier et al. 2009).

Takahashi et al. (2005) used the cat as animal model in experimental periodontitis. In this research, ankylosis at root level during periodontal healing was studied in class III furcation defects surgically created at the level of the premolars.

21.15 Animal Models of Periodontitis-Associated Systemic Disease

Obesity is a major public health concern worldwide in developed and developing countries. Several rat model studies suggested that obesity is a risk factor for several chronic diseases, including periodontitis (Endo et al. 2010; Ekuni et al. 2010; Tomofuji et al. 2009b; Watanabe et al. 2008). Obese Zucker rats, which bear an autosomal recessive mutation of the leptin receptor gene, were used because it is well known that they show chronic hyperphagia resulting in morbid obesity (Zucker 1965; Kasiske et al. 1992). **The Zucker fatty rat (ZFR)** is a well-established model of prediabetes and characterized by hyperphagia, hyperlipidemia, obesity, insulin resistance, hyperinsulinemia and impaired glucose tolerance (Janssen et al. 1999). The mechanisms by which obesity augments periodontitis are not completely understood. Obesity affects host immunity (Martí et al. 2001). The increase in systemic low-grade inflammation induced by periodontitis may alter the effects of obesity on the production of inflammatory molecules, including C-reactive protein, interleukin-6 and tumour

necrosis factor- α , in the liver and white adipose tissue (Endo et al. 2010). Perlstein and Bissada (1977) evaluated the extent to which obesity and/or hypertension may modify the response of rats' periodontium to chronic gingival irritation. Histopathologic evaluation of the periodontal structure showed both hyperplasia and hypertrophy of the walls of blood vessels supplying the periodontium in the hypertensive and obese-hypertensive animals. The results also indicated that obesity significantly contributed to the severity of periodontal disease. The obese-hypertensive rats showed the most severe periodontal response to local irritation.

The development of obesity generally precedes the onset of **insulin resistance** in a person with the metabolic syndrome. **KK-Ay mice**, established as a type 2 diabetes mouse model accompanied by obesity (Nishimura 1969), are considered to be an appropriate mouse model of the syndrome because the mice develop insulin resistance and cardiovascular risk factors such as hypertension, hyperlipemia and glucose tolerance (Iwatsuka et al. 1970; Aizawa-Abe et al. 2000; Ohnishi et al. 2009). Reactive oxygen species, such as hydrogen peroxide, are responsible for the alveolar bone loss accompanied by decreased endothelial nitric oxide synthase expression in KK-Ay mice. Therefore, a working hypothesis was proposed that the generation of oxidative stress is an underlying systemic condition that enhances alveolar bone loss in periodontitis occurring as a complication of diabetes (Ohnishi et al. 2009).

Doxey et al. (1998) had used a rat model to study the effects of **diabetes** on periodontal disease and to characterize the different gingival cytokine profiles in response to *P. gingivalis*-induced periodontal disease. Diabetes was induced chemically by **streptozotocin administration to Sprague–Dawley rats**. A significant increase in the levels of platelet-derived growth factor B (PDGF-B) and IL-1 β was observed in the nondiabetic periodontal disease model. However, these increases did not occur in the diabetes model and were prevented in the model with periodontitis superimposed on diabetes. These results demonstrate that the normal host

response to periodontitis is altered during diabetes by the prevention of periodontitis-induced increases in IL-1 β and PDGFB. This is contrary to the observations in humans; thus, the implications of the use of this model for further studies are uncertain (Genco et al. 1998b). **Zucker diabetic fatty (ZDF) rats** are a widely used, well-characterized model of obesity and to type 2 diabetes mellitus (T2DM) in diabetes research. ZDF rats are a substrain of Zucker fatty rats (ZFR, *falfa*) and have additional abnormalities that predispose them to T2DM. Unlike ZFR rats, which exhibit hyperinsulinemia, insulin resistance and glucose intolerance but do not develop T2DM, ZDF rats develop T2DM. ZFR and ZDF rats have a point mutation in the leptin receptor (*falfa*) that leads to impairment of leptin signalling, whereas the heterozygote for the leptin receptor mutation (*fal+*) is a lean control (Zucker lean control). Leptin has been implicated in the signalling network that modifies hunger and satiety; thus, ZDF rats exhibit hyperphagia. The male ZDF rat develops an age-dependent diabetes phenotype, with the onset of obesity at 5 weeks of age accompanied by a metabolic state of prediabetes with hyperinsulinemia and insulin resistance. Because of this rapid onset, male ZDF rats are not ideal to study the temporal development of IR. In contrast, female ZDF rats develop T2DM only after consuming a high-fat diet. Thus, female ZDF rats may be considered more useful in studying the development of insulin resistance and T2DM, primarily because of the lower onset and also because they become insulin resistant upon consuming a high-fat diet, which models the disease progression in most humans in Western countries. Thus, female ZDF rats are an excellent model for prediabetes and T2DM in obese humans and are ideal to investigate the effect of periodontitis on the onset of insulin resistance and T2DM (Watanabe et al. 2008 and reference therein).

Thus, the above-mentioned diabetes models seem to mimic well the increased susceptibility and progression of periodontitis seen in diabetic patients. In general, genetically diabetic rodent models and type 2 diabetes models have been investigated less. To our knowledge, there are still unexplored diabetes models to be investigated for

Table 21.11 Experimental periodontitis in humanized non-obese diabetic (NOD)/severe combined immunodeficient (SCID) gavage model (Graves et al. 2008) (Reprinted with permission from John Wiley & Sons, Inc.)

T cells significantly contribute to periodontal bone loss and tissue inflammation	Teng et al. (1999, 2000), Teng (2002b)
The cytokine, RANKL plays an important role in periodontal bone loss and its production is associated with both Th1 and Th2 responses during periodontal pathogenesis	Teng et al. (2000), Teng (2002b), Mahamed et al. (2005)
OPG administration significantly abolishes alveolar bone destruction <i>in vivo</i> ; interferon- γ (IFN- γ) treatment increases periodontal bone loss while treatment with interleukin 10(IL-10) reduces alveolar bone loss	Teng et al. (2005), Zhang and Teng (2006)
The antigen CagE found on <i>Aggregatibacter actinomycetemcomitans</i> , which is pro-apoptotic, plays a significant role in periodontal tissue and bone loss	Teng and Hu (2003), Teng (2003), Teng and Zhang (2005)
CD11c1 dendritic cells can participate in modulating inflammation-induced bone loss by acting as osteoclast precursors	Alnaeeli et al. (2006, 2007)

the susceptibility to periodontitis, such as models of prediabetes and type 2 diabetes models, that is, *ob/ob* mice and OLETF and the model of type 1 diabetes, that is, the BB rat. Obese (*ob/ob*) mice are leptin deficient as a result of a mutation in the obese (*ob*) gene, which is why they present mild hyperglycemia and become prediabetic. Male Otsuka Long–Evans Tokushima fatty rats (OLETFs) develop type 2 diabetes associated with obesity based on multiple recessive genes. Impaired glucose tolerance occurs as early as 8 weeks of age and progresses to hyperglycemia after 18 weeks. The Bio-Breeding (BB) rat is a model of genetic type 1 diabetes usually presenting an abrupt onset of the disease at the age of 8–12 weeks that is characterized by the classic signs, that is, glucosuria, hyperglycemia, ketonuria, hypoinsulinemia and weight loss. Moreover, comparison of different methods of periodontitis induction has not been performed in diabetic rodents. Research in this area is warranted; it might assist in choosing the most appropriate methods and elucidating the contribution of the use of periodontal bacteria (Pontes Andersen et al. 2007).

Because most human periodontal pathogens are not habitants in mice and laboratory mouse strains, rendering an *in vivo* study of human periodontal disease difficult, humanization of mouse models provides an excellent opportunity to investigate the underlying processes or mechanisms regarding specific aspects of immune–bacteria interactions related to aspects of periodontal pathogenesis. Despite being considered exogenous pathogen(s) in mice under study, the second-

ary (or recall) immune responses and periodontal tissue inflammation against *A. actinomycetemcomitans* infection can be readily observed and measured in a humanized **immunodeficient NOD/SCID model** using the oral gavage approach (see **Table 21.11**) (Graves et al. 2008).

Potential and future uses of this model

1. To investigate mechanisms by which diabetes exacerbates periodontal bone loss
2. To identify susceptibility genes that lead to enhancing periodontal bone loss
3. To examine the role of leucocyte subsets in promoting periodontal inflammation and alveolar bone loss

Recent animal studies (Endo et al. 2010; Tomofuji et al. 2007, 2008, 2009a, b; Hyvärinen et al. 2009) revealed that periodontitis and/or periodontal pathogens could induce **hepatic inflammation** (**Fig. 21.69**). In a rat model, application of lipopolysaccharide and proteases to the gingival sulcus augmented the effect of a high-cholesterol diet on steatosis, inflammation and oxidative damage in the liver (**Fig. 21.70**) (Yamamoto et al. 2010).

The mechanistic aspect of the possible association of periodontal disease with **pregnancy complications** has been explored in several experimental animal models. In most models, periodontal bacteria (*P. gingivalis* or *Campylobacter rectus*) are injected in a small chamber that previously had been implanted subcutaneously in the pregnant small animals (hamsters, mice, rabbits) (Bobetsis et al. 2006; Arce et al. 2009; Lin et al. 2003), while in baboons, the ligature-induced periodontitis

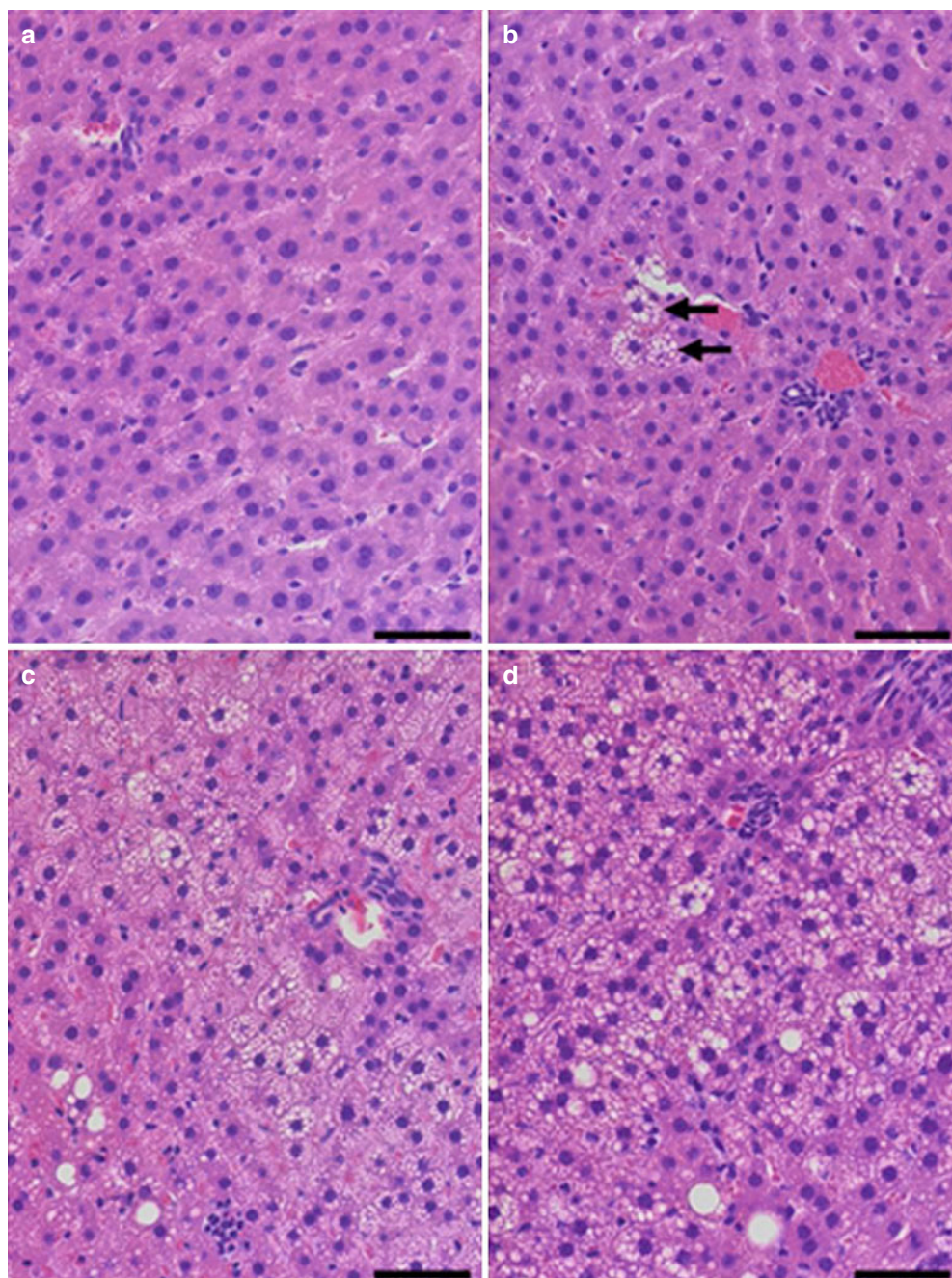


Fig. 21.69 Male Wistar rats were divided into four groups. During the experimental period of 8 weeks, two groups were fed an ethanol-containing liquid diet, and two groups were on a pair-fed control diet. Four weeks prior to the end of the experimental period, one group from each dietary treatment was ligated to induce periodontitis, while the other group was left unligated. Liver specimens from rats in the Control, Periodontitis, Ethanol and Combination groups stained with haematoxylin and

eosin. Few pathological changes were observed in the Control group (a). The Periodontitis group showed slight fatty change characterized by small droplets (arrows) in the hepatocytes (b). The fatty changes of Ethanol (c) and Combination (d) groups were severer than that of the Periodontitis group. The degree of steatosis and inflammation was greater in the Combination group (d) than the other groups. Bar=25 μ m (Tomofuji et al. 2008. Reprinted with permission from Elsevier)

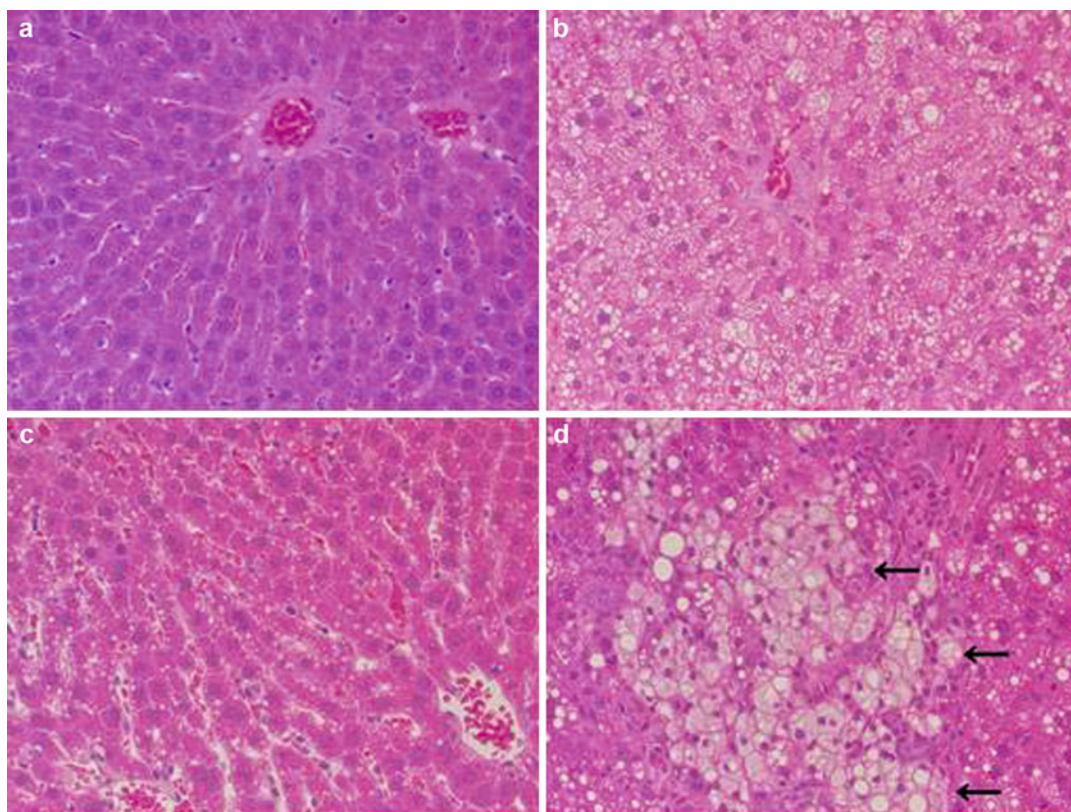


Fig. 21.70 Liver from rats in the control, cholesterol, periodontitis and combination groups stained with hematoxylin and eosin. Few pathological changes were observed in the control group (a). The periodontitis group showed slight fatty change, characterized by small droplets in the hepatocytes (c). The fatty changes of cholesterol

(b) and combination groups (d) were more severe than those of the periodontitis group (c). Foci of hepatocytes containing fat vacuoles (arrows) are more common in the combination group (d) than any other groups. Scale bar represents 50 μ m (Yamamoto et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

model was used (Cappelli et al. 2009; Ebersole et al. 2010). The results of these studies reveal that maternal infection with periodontal pathogens has a deleterious effect on foetal growth and viability. Specifically, both *P. gingivalis* and *C. rectus* have the capacity to disseminate from the subcutaneous chamber not only towards maternal organs (liver, uterus) but most importantly towards placental and foetal tissues. This translocation is accompanied by an increase in inflammatory mediators in the placenta. Moreover, the infection with periodontal pathogens induces a significant alteration in the architecture of the placenta, especially in areas that are critical for the exchange of nutrients between the mother and the foetus. Furthermore, maternal exposure to *P. gingivalis* or *C. rectus* results in a

decrease in the size of the foetuses (preterm deliveries do not occur in mice). The reduced size of the foetuses is not the only complication, since the newborns demonstrate a higher risk of experiencing perinatal death, similarly to PLBW human infants. Finally, pups that survive the perinatal period appear to have an increase in inflammatory cytokines (interferon- γ) in the brain tissues along with ultrastructural alterations in the hippocampal region of the brain. Interestingly, these changes in the neonatal brain occur in a manner analogous to the effect of maternal infection on white-matter damage seen in humans. Taken together, these findings suggest that the threat of maternal infection with periodontal pathogens during pregnancy may not be limited to the duration of

Fig. 21.71 Macroscopic appearance of the paws and gingival tissues. (a) Experimental arthritis (EA) only front paw; (b) Experimental arthritis (EA) only rear paw; (c) periodontitis and EA front paw; (d) periodontitis and EA rear paw; (e) control gingiva; (f) periodontitis gingiva. Paws were imaged at day 8 post arthritis induction (Cantley et al. 2011. Reprinted with permission from John Wiley & Sons, Inc.)



gestation but also may affect perinatal neurological growth and development (Bobetsis et al. 2006). Collins et al. (1994a) found that LPS from *P. gingivalis* significantly reduced the foetal weight, suggesting that maternal exposure to *P. gingivalis* LPS can have harmful effects on the developing foetus. It has also been shown that foetal growth retardation and embryoletality are significantly correlated with increases in PGE2 and TNF- α occurring after *P. gingivalis* challenge (Collins et al. 1994b).

A growing number of epidemiologic, serologic and animal model studies provided evidence that *P. gingivalis*, the major etiological agent of periodontal disease, might be involved in the

onset and progression of **rheumatoid arthritis** (Figs. 21.71 and 21.72) (Liao et al. 2009; Mikuls et al. 2009; Cantley et al. 2011; Bartold et al. 2005, 2010; Ebersole et al. 2002; Ramamurthy et al. 2005).

Various animal studies provide evidence of an association between periodontal disease and **depression** (Soletti et al. 2009; Breivik et al. 2006) or between **stress** and inflammatory periodontal disease (da Monteiro Silva et al. 1995; Breivik et al. 2000a, b; Nakajima et al. 2006; Peruzzo et al. 2008; Semenoff Segundo et al. 2010; Okada et al. 2010). Breivik et al. (2006) tested a model in which the extent of periodonti-

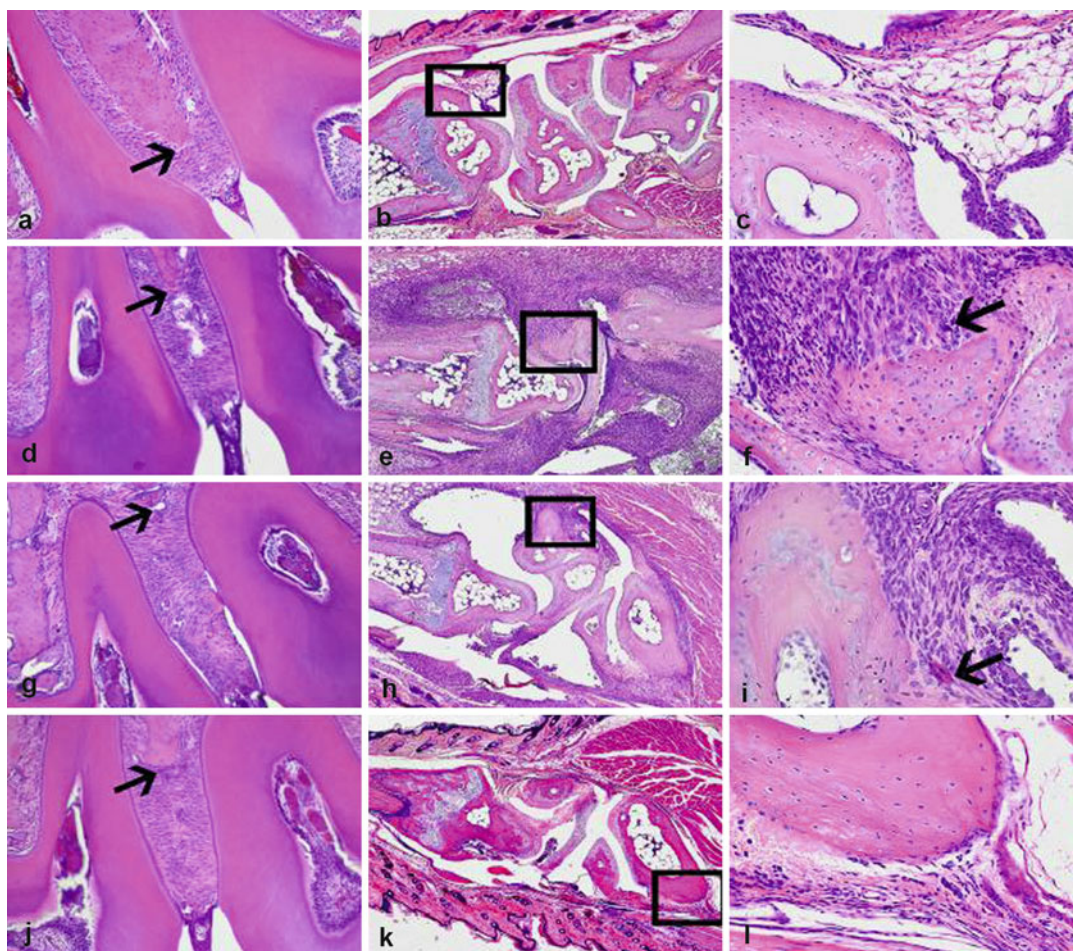


Fig. 21.72 Histological appearance of periodontal tissues (a, d, g, j) at 20× magnification and front paw joints (b, e, h, k) at 4× magnification and (c, f, i, l) at 20× magnification. (a, b, c) Controls (d, e, f) periodontitis and experimental arthritis (EA) (g, h, i) EA only (j, k, l) periodontitis only. The boxes in the second panel highlight the

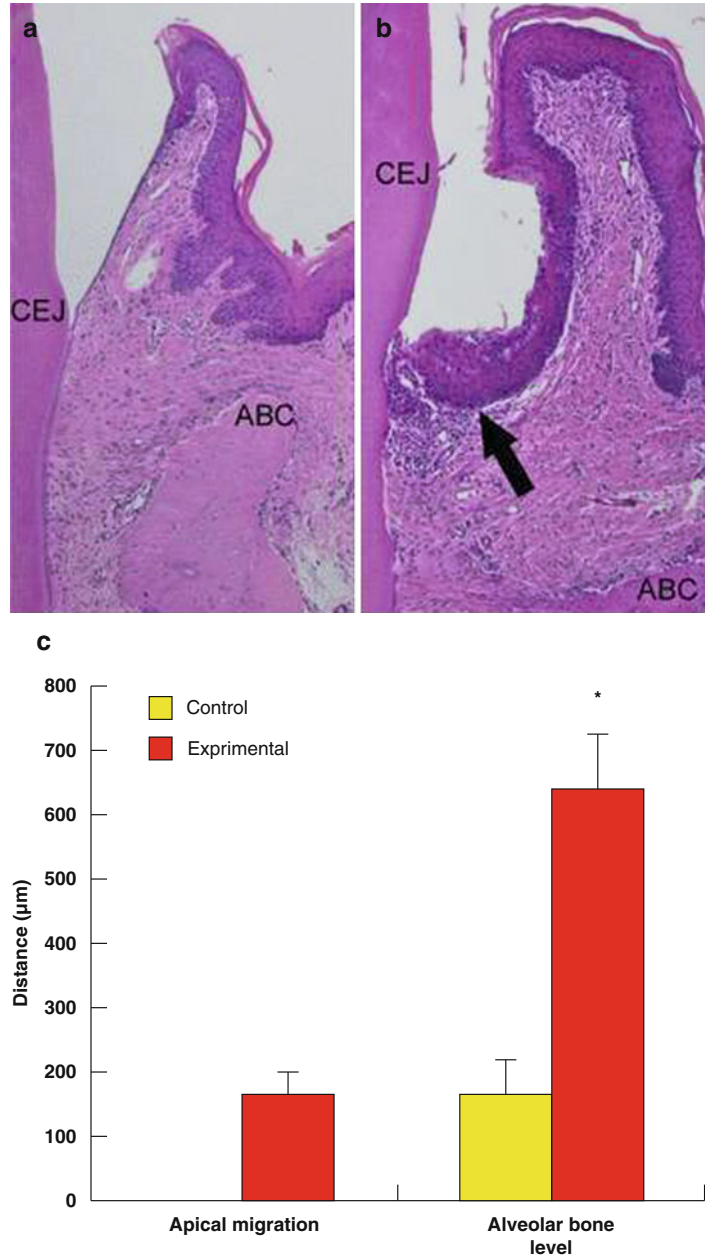
areas imaged at 20× magnification in the third panel. Arrows in the first panel indicate level of alveolar bone crest. In the third panel, arrows indicate the presence of multinucleated osteoclast cells (Cantley et al. 2011). Reprinted with permission from John Wiley & Sons, Inc.)

tis, an inflammatory disease, is predicted from the reactivity of the HPA axis. Briefly, the model implies that elevated plasma levels of corticosterone modulate the immune system by shifting the T-helper balance towards a more Th2 response, which, alternately, increases the sensitivity to periodontitis. Moreover, these data strengthen conclusions from experimental studies in two genetically selected rat lines with extreme differences in their behavioural and neuroendocrine responses to stress, in which were found significant differences in susceptibility to peri-

odontal disease (Breivik et al. 2000b). It has also been shown that manipulations in maternal environment reverse periodontitis in genetically predisposed rats (Sluyter et al. 2002).

There are recent reports of the association between cardiovascular disease and periodontal disease, and it will be critical to develop animals that will enable investigators to determine the pathogenic potential of *P. gingivalis* in models of cardiovascular disease (Genco et al. 1998b). A ligature-induced periodontitis dog model revealed that periodontitis led to inflammatory responses in

Fig. 21.73 Pathological changes in rat periodontal tissue. While the control group showed no pathological changes (a), the experimental group showed alveolar bone resorption, apical migration of junctional epithelium (*arrow*) and inflammatory cell infiltration in the connective tissue adjacent to the junctional epithelium (b) (hematoxylin and eosin staining). Distances between the cemento-enamel junction and the apical portion of the junctional epithelium (apical migration), and between the cemento-enamel junction and the alveolar bone crest (alveolar bone level) were larger in the experimental group than in the control group (c) (mean $\pm$ standard deviation). *ABC* alveolar bone crest, *CEJ* cemento-enamel junction. Scale bar = 200 μ m. $*P < 0.001$, compared with the control group (Mann–Whitney *U*-test) (Ekuni et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)

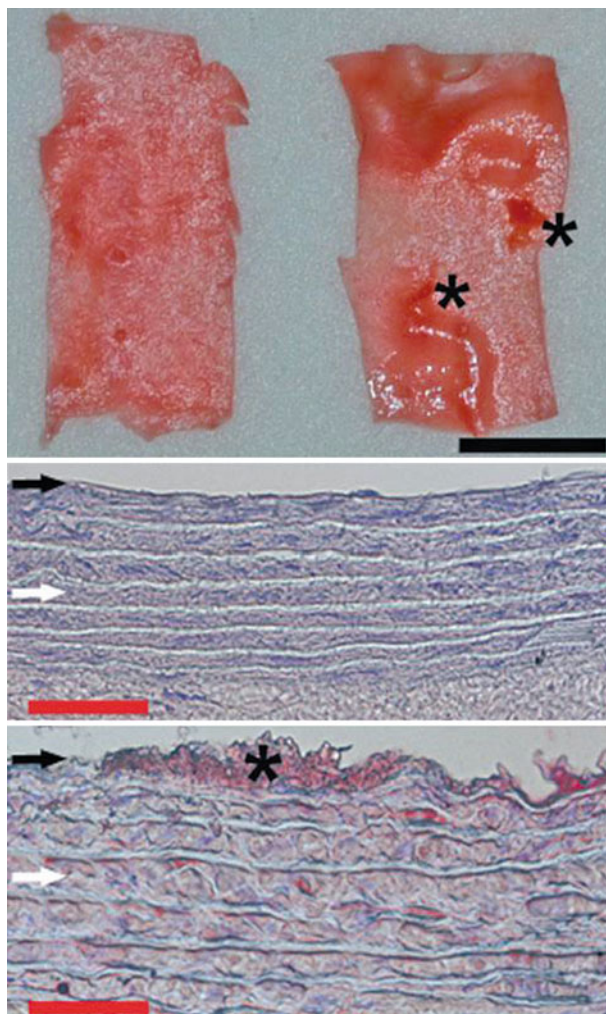


the atrial myocardium, which disturbed the structural and electrophysiologic properties of the atrium and facilitated **atrial fibrillation** (Yu et al. 2010).

Although DNA from oral bacteria has been found in atherosclerotic plaque in animal experimental model (Zhang et al. 2008), the contribution of these bacteria to plaque formation remains unknown (Hayashi et al. 2010). In a ligature-

induced periodontitis rat model, increased lipid peroxidation was found in serum and aorta as well as in periodontal tissue. Atherosclerosis-related gene expression and histological changes were also stimulated. Periodontitis-induced lipid peroxidation in the aorta may be involved in the early stage of atherosclerosis (Figs. 21.73 and 21.74) (Ekuni et al. 2009).

Fig. 21.74 Representative results of lipid deposition in the descending aorta stained with *oil red O*. Lipid deposition (*asterisk*) was observed in the experimental group (B, D) but not in the control group (A, C). *Black arrows* intima; *white arrows* media; *black scale bar*=1 mm; *red scale bar*=50 μ m (Ekuni et al. 2009. Reprinted with permission from John Wiley & Sons, Inc.)



Studies modelling the **periodontitis/atherosclerosis** associations used genetically modified ***ApoE*-deficient mice**, a mouse model prone to accelerated atherosclerosis on a high-fat, high-cholesterol diet ('Western diet') (Kebschull et al. 2010). The induction of periodontal disease by intravenous injection of *P. gingivalis* (Li et al. 2002b), *P. gingivalis* LPS (Gitlin and Loftin 2009) or repeated oral/anal bacterial applications (Lalla et al. 2003) resulted in enhanced atherosclerosis in infected animals when compared with uninfected controls (Fig. 21.75) (Kebschull et al. 2010). However, only invasive *P. gingivalis*, as compared with a fimbriae-deficient mutant, was able to induce periodontitis, increase atherogenesis and trigger up-regulation of Toll-like

receptors TLR2 and TLR4 (Gibson et al. 2004). Due to the used animal model, all the obtained findings may have limited relevance to the human situation (Kebschull et al. 2010). In contrast, wild-type mice were shown to be relatively resistant to a plethora of atherogenic stimuli and did not develop atherosclerosis within reasonable time periods (Graves et al. 2008; Kebschull et al. 2010). However, in other animal models involving pigs, *P. gingivalis* was shown to induce rapid atherogenesis, even in normo-cholesterolemic situations (Fig. 21.76) (Brodala et al. 2005). Nevertheless, cost-related issues and, more importantly, the possibility to generate knock-out or knock-in animals make mouse models

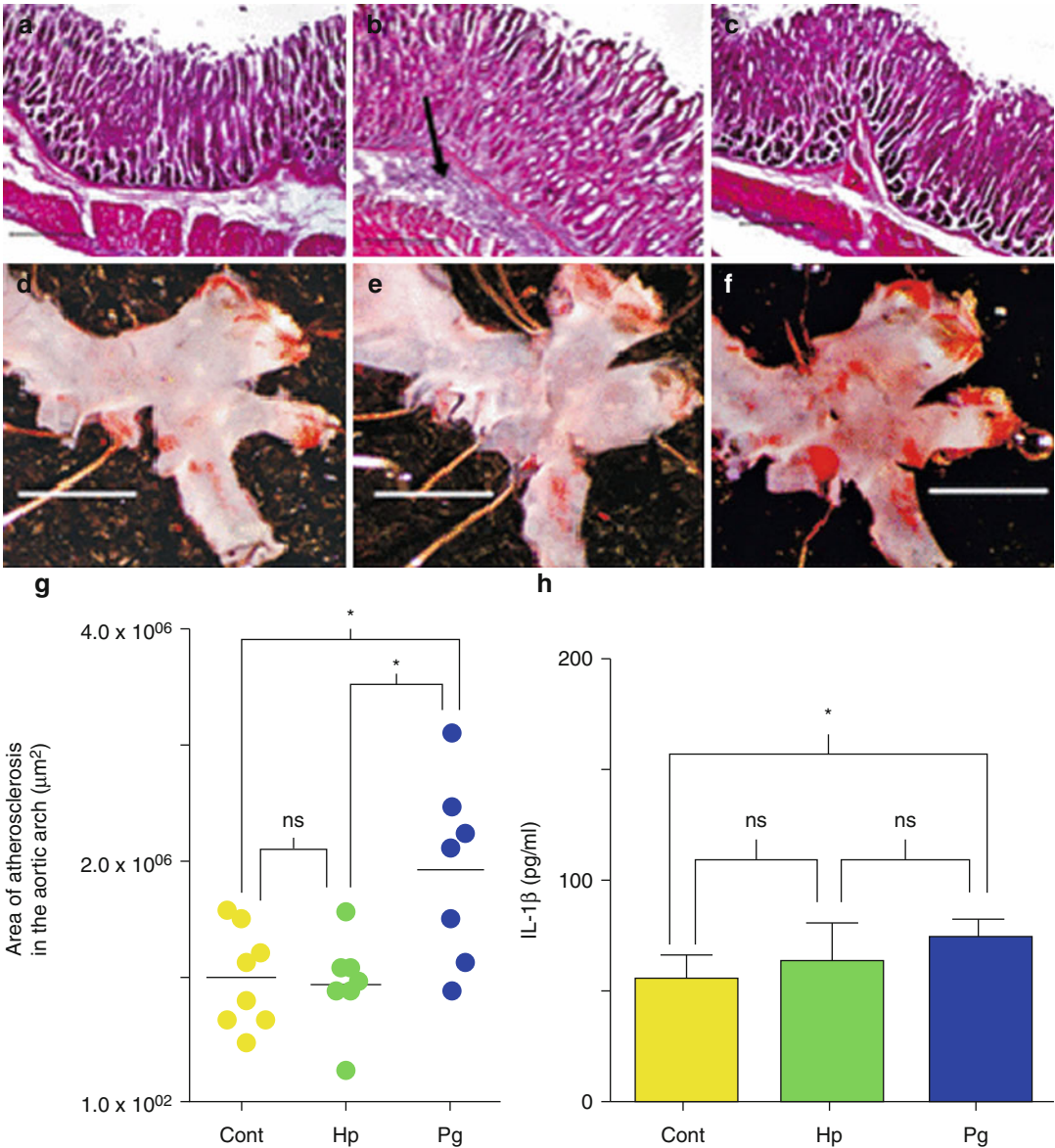


Fig. 21.75 Chronic infection is established in apolipoprotein A-deficient ($\text{ApoE}^{-/-}$) mice challenged with *Porphyromonas gingivalis* and *Helicobacter pylori*, yet only *P. gingivalis*-infected mice present with acceleration of atherosclerosis. Stomach tissue from representative uninfected (a), *H. pylori*-infected (b) and *P. gingivalis*-infected (c) mice reveal that *H. pylori*-infected mice ($n=8/\text{group}$) present with submucosal cellular infiltrate (arrow); representative micrographs of Sudan IV staining for lipids (red) on the intimal surface of the aortic arch region from uninfected (d), *H. pylori*-infected (e) and *P. gingivalis*-

infected (f) mice show increased staining only in *P. gingivalis*-challenged mice, which was confirmed by morphometric analysis (g). Horizontal line denotes the group median and black circle represents an individual mouse. Serum levels of the proinflammatory cytokine interleukin-1 β (IL-1 β) as measured by enzyme-linked immunosorbent assay (h); * $P<0.05$ by analysis of variance with Dunn's multiple comparisons; ns not significantly different; scale bars in a–c=25 μm , and d–f=1 mm (Hayashi et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

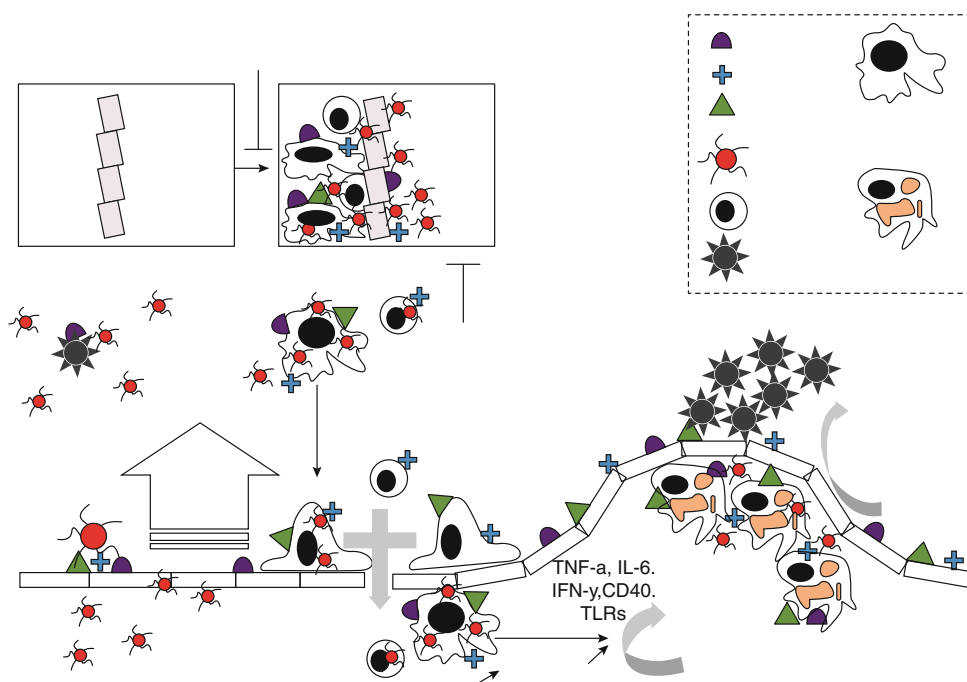


Fig. 21.76 Model for *Porphyromonas gingivalis* systemic infection and acceleration of atherosclerosis. (1) Platelets, monocytes and dendritic cells are depicted in circulation. (2) Monocytes and dendritic cells are depicted in the oral mucosal lesion. Vaccination prevents oral colonization whereas antibiotics decrease the bacterial burden. Platelets, monocytes, dendritic cells, and foam cells are

depicted at the atheroma. *CAMs* cell adhesion molecules, *IFN* interferon, *IL* interleukin, *M*/ macrophages, *LDL* low-density lipoprotein, *PD* periodontal diseases, *SE* sulcular epithelium, *SRs* scavenger receptors, *TNF* tumour necrosis factor, *VE* vascular endothelium (Hayashi et al. 2010. Reprinted with permission from John Wiley & Sons, Inc.)

extremely useful in the study of specific molecular targets of atherogenesis (Kebschull et al. 2010).

21.16 Improving the Reporting of Animal Experiments: The ARRIVE Guidelines

Failure to describe research methods and to report results appropriately therefore has potential scientific, ethical and economic implications for the entire research process and the reputation of those involved in it. This is particularly true for animal research, one of the most controversial areas of science (Kilkenny et al. 2010). Improved reporting of study details will maximize the availability and utility of the information gained from every animal and every experiment, preventing unnecessary animal use in the future. To address this, we led an initiative to produce guidelines for

reporting animal research (Kilkenny et al. 2010). The guidelines, referred to as **ARRIVE (Animals in Research: Reporting In Vivo Experiments)**, have been developed using the CONSORT Statement as their foundation (Schulz et al. 2010; Moher et al. 2011).

The ARRIVE guidelines consist of a checklist of 20 items describing the minimum information that all scientific publications reporting research using animals should include, such as the number and specific characteristics of animals used (including species, strain, sex and genetic background); details of housing and husbandry; and the experimental, statistical and analytical methods (including details of methods used to reduce bias such as randomization and blinding) (Kilkenny et al. 2010).

The ARRIVE guidelines (see **Table 21.12**) can be applied to any area of bioscience research using laboratory animals, and the inherent principles

Table 21.12 Animal research: reporting *in vivo* experiments: the ARRIVE guidelines Kilkenney et al. (2010). Reprinted under the terms of the Creative Commons Attribution License

	Item	Recommendation
Title	1	Provide as accurate and concise a description of the content of the article as possible
Abstract	2	Provide an accurate summary of the background, research objectives (including details of the species or strain of animal used), key methods, principal findings and conclusions of the study
Introduction		
Background	3	(a) Include sufficient scientific background (including relevant references to previous work) to understand the motivation and context for the study and explain the experimental approach and rationale (b) Explain how and why the animal species and model being used can address the scientific objectives and, where appropriate, the study's relevance to human biology
Objectives	4	Clearly describe the primary and any secondary objectives of the study or specific hypotheses being tested
Methods		
Ethical statement	5	Indicate the nature of the ethical review permissions, relevant licences (e.g. Animal [Scientific Procedures] Act 1986) and national or institutional guidelines for the care and use of animals, that cover the research
Study design	6	For each experiment, give brief details of the study design, including: (a) The number of experimental and control groups (b) Any steps taken to minimize the effects of subjective bias when allocating animals to treatment (e.g. randomization procedure) and when assessing results (e.g. if done, describe who was blinded and when) (c) The experimental unit (e.g. a single animal, group or cage of animals). A time line diagram or flowchart can be useful to illustrate how complex study designs were carried out
Experimental procedures	7	For each experiment and each experimental group, including controls, provide precise details of all procedures carried out. For example: (a) How (e.g. drug formulation and dose, site and route of administration, anaesthesia and analgesia used [including monitoring], surgical procedure, method of euthanasia). Provide details of any specialist equipment used, including supplier(s) (b) When (e.g. time of day) (c) Where (e.g. home cage, laboratory, water maze) (d) Why (e.g. rationale for choice of specific anaesthetic, route of administration, drug dose used)
Experimental animals	8	(a) Provide details of the animals used, including species, strain, sex, developmental stage (e.g. mean or median age plus age range) and weight (e.g. mean or median weight plus weight range) (b) Provide further relevant information such as the source of animals, international strain nomenclature, genetic modification status (e.g. knock-out or transgenic), genotype, health/immune status, drug- or test naïve, previous procedures, etc.
Housing and husbandry	9	Provide details of: (a) Housing (e.g. type of facility, e.g. specific pathogen free (SPF); type of cage or housing; bedding material; number of cage companions; tank shape and material etc. for fish) (b) Husbandry conditions (e.g. breeding program, light/dark cycle, temperature, quality of water etc. for fish, type of food, access to food and water, environmental enrichment) (c) Welfare-related assessments and interventions that were carried out before, during or after the experiment
Sample size	10	(a) Specify the total number of animals used in each experiment and the number of animals in each experimental group (b) Explain how the number of animals was decided. Provide details of any sample size calculation used (c) Indicate the number of independent replications of each experiment, if relevant

Table 21.12 (continued)

Allocating animals to experimental groups	11	(a) Give full details of how animals were allocated to experimental groups, including randomization or matching if done (b) Describe the order in which the animals in the different experimental groups were treated and assessed
Experimental outcomes	12	Clearly define the primary and secondary experimental outcomes assessed (e.g. cell death, molecular markers, behavioural changes)
Statistical methods	13	(a) Provide details of the statistical methods used for each analysis (b) Specify the unit of analysis for each dataset (e.g. single animal, group of animals, single neuron) (c) Describe any methods used to assess whether the data met the assumptions of the statistical approach
Results		
Baseline data	14	For each experimental group, report relevant characteristics and health status of animals (e.g. weight, microbiological status and drug- or test-naïve) before treatment or testing (this information can often be tabulated)
Numbers analyzed	15	(a) Report the number of animals in each group included in each analysis. Report absolute numbers (e.g. 10/20, not 50%) (b) If any animals or data were not included in the analysis, explain why
Outcomes and estimation	16	Report the results for each analysis carried out, with a measure of precision (e.g. standard error or confidence interval)
Adverse events	17	(a) Give details of all important adverse events in each experimental group (b) Describe any modifications to the experimental protocols made to reduce adverse events
Discussion		
Interpretation/scientific implications	18	(a) Interpret the results, taking into account the study objectives and hypotheses, current theory and other relevant studies in the literature (b) Comment on the study limitations including any potential sources of bias, any limitations of the animal model and the imprecision associated with the results (c) Describe any implications of your experimental methods or findings for the replacement, refinement or reduction (the 3Rs) of the use of animals in research
Generalisability/translation	19	Comment on whether, and how, the findings of this study are likely to translate to other species or systems, including any relevance to human biology
Funding	20	List all funding sources (including grant number) and the role of the funder(s) in the study

apply not only to reporting comparative experiments but also to other study designs. Laboratory animal refers to any species of animal undergoing an experimental procedure in a research laboratory or formal test setting. The guidelines are not intended to be mandatory or absolutely prescriptive, nor to standardize or formalize the structure of reporting. Rather they provide a checklist that can be used to guide authors preparing manuscripts for publication and by those involved in peer review for quality assurance, to ensure completeness and transparency (Kilkenny et al. 2010).

The largest and most comprehensive review of published animal research undertaken to date, to

our knowledge, has highlighted serious omissions in the way research using animals is reported. Detailed information was collected from 271 publications, about the objective or hypothesis of the study, the number, sex, age and/or weight of animals used and experimental and statistical methods. Only 59% of the studies stated the hypothesis or objective of the study and the number and characteristics of the animals used. Most of the papers surveyed did not use randomization (87%) or blinding (86%) to reduce bias in animal selection and outcome assessment. Only 70% of the publications that used statistical methods described their methods and presented the results with a measure of error or variability.

Fig. 21.77 Review authors' judgement, on the basis of each criterion, of the risk of bias arising, presented as a percentage of the studies published from 2008 to 2010 (Faggion et al. 2011). Reprinted with permission from John Wiley & Sons, Inc.)

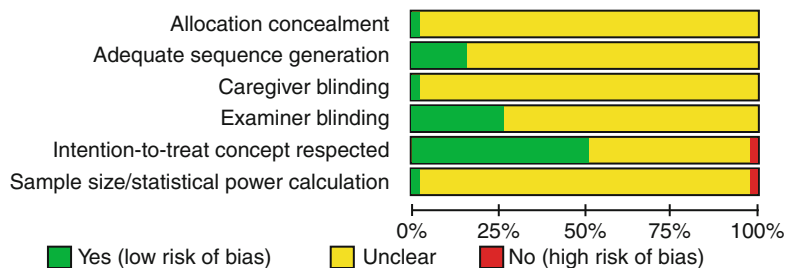
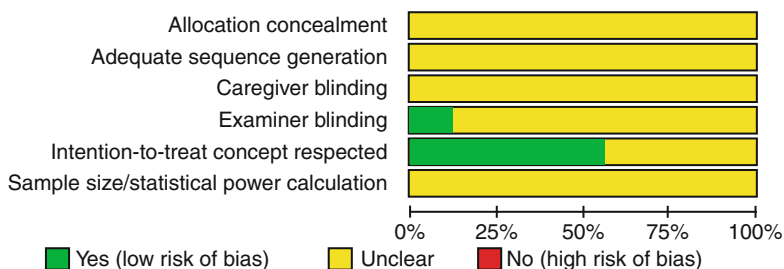


Fig. 21.78 Review authors' judgement, on the basis of each criterion, of the risk of bias arising, presented as a percentage of the studies published from 1998 to 2000 (Faggion et al. 2011). Reprinted with permission from John Wiley & Sons, Inc.)



This survey has identified a number of issues that need to be addressed in order to improve experimental design and reporting in publications describing research using animals (Kilkenny et al. 2009).

Recently, Faggion et al. (2011) assessed the risk of bias of studies in which animal models were used for investigating regenerative therapies for periodontal and peri-implant defects and investigated changes in risk of bias by comparing samples drawn from two different publication periods. The PubMed and LILACS electronic databases, independently and in duplicate, were searched for randomized and controlled trials published from 1998 to 2000 and from 2008 to 2010. The risk of bias was assessed with reference to a six-item checklist based on the Cochrane Collaboration's tool for assessing the risk of bias and information about formal sample-size calculation. Of the 749 items assessed for 107 papers (i.e. seven items for each paper), 581 (78%) were at unclear risk of bias. Domains related to the randomization process were almost never reported. Allocation concealment method and adequate sequence generation were never reported in the sample of studies published from 1998 to 2000. In the 2008–2010 sample, allocation concealment method and adequate sequence

generation were reported once (1.5% of the sample) and ten times (15% of the sample), respectively. Blinding of caregivers was never reported in the sample of 1998–2000, whereas in the other sample, this topic was reported only once (1.5% of the sample). Examiner blinding was reported five times (12% of the sample) and 17 times (26% of the sample) in the 1998–2000 and 2008–2010 samples, respectively. Only one study in the two samples reported to have performed sample-size/statistical power calculation. **Figures 21.77** and **21.78** illustrate, graphically, differences between results from assessment of the methodology of studies published from 1998 to 2000 and those published from 2008 to 2010 (Faggion et al. 2011).

21.17 Methodological Quality of Systematic Reviews of Animal Studies

Systematic reviews can serve as a tool in translation of basic life sciences research from laboratory to human research and health care. It has been showed that systematic reviews of animal studies often lacked methodological features such as specification of a testable hypothesis (30%);

literature search without language restriction (26.6%); assessment of publication bias (16.6%), study validity (50%) and heterogeneity (33.3%); and meta-analysis for quantitative synthesis (40%). Compared to reviews of bench studies, they were less prone to bias as they specified the question (96.6% vs. 80%, $P=0.04$), searched multiple databases (60% vs. 26.6%, $P=0.01$), assessed study quality (50% vs. 20%, $P=0.01$) and explored heterogeneity (33.3% vs. 2.2%, $P=0.001$) more often. There seems to be a gradient of frequency of methodological weaknesses among reviews: Attempted systematic reviews of whole-animal research tend to be better than those of bench studies, though compared to systematic reviews of human clinical trials they are apparently poorer (Mignini and Khan 2006).

When the methodological quality of systematic reviews and meta-analyses of animal studies in dentistry was assessed through using the AMSTAR checklist and of 444 studies retrieved, 54 systematic reviews were selected after full-text assessment, only two studies were scored as high quality, and 17 and 35 studies were scored as medium and low quality, respectively (Faggion et al. 2012). These findings suggested that there are substantial benefits to be obtained from further improvement of the methodological quality of systematic reviews on animal experiments in dentistry. Use of validated checklists such as AMSTAR can improve the design, conduct, and reporting of systematic reviews and meta-analysis. Well-designed animal experiments need to be considered an essential ingredient of robust systematic reviews. Systematic reviews are not able to improve the data obtained from poorly designed studies. However, systematic reviews may provide estimates which can be generalized beyond the specific conditions of a particular study. Moreover, systematic reviews may improve the efficiency in research by contributing to a more comprehensive use of data generated by single animal experiments. In this context, high-quality systematic reviews might allow a reduction in the need for further animal experiments by relying on and emphasizing the value of already existing information (Faggion et al. 2012).

21.18 Comparison of Treatment Effects Between Animal Experiments and Clinical Trials

Clinicians and the public often consider it axiomatic that animal research has contributed to the treatment of human disease, yet little evidence is available to support this view. Few methods exist for evaluating the clinical relevance or importance of basic animal research, and so, its clinical (as distinct from scientific) contribution remains uncertain. Anecdotal evidence or unsupported claims are often used as justification—for example, statements that the need for animal research is ‘self-evident’ or that ‘animal experimentation is a valuable research method which has proved itself over time’. Such statements are an inadequate form of evidence for such a controversial area of research. Moreover, if animal experiments fail to inform medical research, or if the quality of the experiments is so poor as to render the findings inconclusive, the research will have been conducted unnecessarily (Table 21.13) (Pound et al. 2004).

In medical research, Hackam and Redelmeier (2006) conducted a systematic review to determine how often highly cited animal studies translate into successful human research. Of the animal studies, 37% (95% confidence interval, 26–48%) were replicated in human randomized trials, 18% were contradicted by randomized trials, and 45% remain untested. Median time to replication was 7 years (range, 1–15 years). Global methodology score did not predict translation in unadjusted analyses (odds ratio [OR]=1.28 per 10% higher score; 95% CI: 0.97–1.69) or in analyses adjusted for citation rate and length of time available for human replication (OR=1.27; 95% CI: 0.96–1.69). Animal studies incorporating dose–response gradients were more likely to translate to humans (OR=3.3; 95% CI: 1.1–10.1). Other quality criteria, type of therapy, type of disease, species, journal, citation rate, length of follow-up and year of publication did not predict subsequent translation. These findings have important implications: First, patients and physicians should remain cautious about extrapolating the findings of prominent animal research to the care of

Table 21.13 Methodological problems of animal experiments (Pound et al. 2004) (Reprinted with permission from BMJ Publishing Group Ltd.)

• Disparate animal species and strains, with a variety of metabolic pathways and drug metabolites, leading to variation in efficacy and toxicity
• Different models for inducing illness or injury with varying similarity to the human condition
• Variations in drug dosing schedules and regimen that are of uncertain relevance to the human condition
• Variability in the way animals are selected for study, methods of randomization, choice of comparison therapy (none, placebo, vehicle) and reporting of loss to follow up
• Small experimental groups with inadequate power, simplistic statistical analysis that does not account for potential confounding and failure to follow intention to treat principles
• Nuances in laboratory technique that may influence results may be neither recognized nor reported—e.g. methods for blinding investigators
• Selection of a variety of outcome measures, which may be disease surrogates or precursors and which are of uncertain relevance to the human clinical condition
• Length of follow up before determination of disease outcome varies and may not correspond to disease latency in humans

human disease. Second, major opportunities for improving study design and methodological quality are available for preclinical research. Finally, poor replication of even high-quality animal studies should be expected by those who conduct clinical research (Hackam and Redelmeier 2006).

Insights into the limitations of animal models, including the extent to which they represent disease in humans, were demonstrated by Perel et al. (2007) who examined the concordance between treatment effects in animal experiments and clinical trials. Animal studies for interventions with unambiguous evidence of a treatment effect (benefit or harm) in clinical trials were included as follows: head injury, antifibrinolytics in haemorrhage, thrombolysis in acute ischaemic stroke, tirilazad in acute ischaemic stroke, antenatal corticosteroids to prevent neonatal respiratory distress syndrome and bisphosphonates to treat osteoporosis. Concordance between animal studies and clinical studies varied for six interventions: three had similar outcomes and three did not. For example, tirilazad was associated with a worse outcome in patients with ischaemic stroke, while in animal models, tirilazad reduced infarct volume by 29% (21–37%) and improved neurobehavioural scores by 48% (29–67%). Bisphosphonates increased bone mineral density in patients with osteoporosis. Also, in animal models, the bisphosphonate alendronate increased bone mineral density compared with placebo by

11.0% for the hip region, by 8.5% for the lumbar spine and 1.7% for the forearms.

In dental research, animal experiments for testing new surgical intervention have been used for decades. For instance, in research on periodontal regeneration, histological evidence of formation of new periodontal ligament and cementum was often first obtained in animal studies, as the opportunity to obtain histological evidence from humans is usually rare (Faggion et al. 2010). Nevertheless, direct translation of research results from animal models to treatment of human diseases may not always be tenable, because anatomical features and physiological systems are different across species (Perel et al. 2007; Faggion et al. 2010). Moreover, the pathogenesis induced in the animal models may not be similar to that in naturally occurring human diseases. There is always a knowledge gap in the translation of results from animal research into treatment of human diseases (Faggion et al. 2010).

A great heterogeneity between human and animal studies in terms of study designs and treatment procedures were demonstrated when the effects of treatment of peri-implant infection between animal and human studies were compared ($P=97.6\%$ for animal studies and 99.9% for human studies). The single-level and multilevel random-effects meta-analysis showed that the difference in PPD reduction [0.31 mm, 95% confidence interval (CI): $-0.27, 0.88$] and in PAL

Table 21.14 Characteristic of selected studies regarding study design and outcome measures (Faggion et al. 2009) (Reprinted with permission from Elsevier)

Characteristics	Human studies	Animal studies
Study design		
Parallel prospective	13	1
Split-mouth	1	17
Case-series	16	5
Outcome		
Clinical or combined	25	10 ^a
Histological, microbiological, radiological or combined (without clinical outcomes)	5	13

^aSix studies had no outcomes comparable with those from human trials

gain (0.21 mm, 95% CI: -0.47, 0.88) between animal and human studies was not statistically significant. The random-effects meta-regression suggested that studies with longer follow-up times revealed greater PPD reduction (0.25 mm/month, 95% CI: 0.14, 0.35). The results from this study demonstrated the limitations to drawing conclusions from the literature on the management of peri-implantitis. Although the results seemed to suggest that the effects of treatment reported in animal studies were greater than those reported in human studies, the lack of homogeneity in study design and data analysis between the animal and human studies means that the direct comparisons might be questionable. Nevertheless, many serious issues in the study designs, statistical analysis, and reports of the results were identified. These issues need to be taken into consideration in the future research on testing the efficacy of peri-implantitis treatments in both animal and human studies in order to improve the quality of research and reduce the heterogeneity between studies.

When assessment of replication of research evidence from animals to humans in studies on peri-implantitis therapy was performed, it was revealed that (1) results from animal experiments on peri-implantitis may not be extrapolated to the clinical setting because of differences across species and the great variability of research procedures and study designs (Table 21.14 and Fig. 21.79) and (2) the power of animal experiments and human trials should

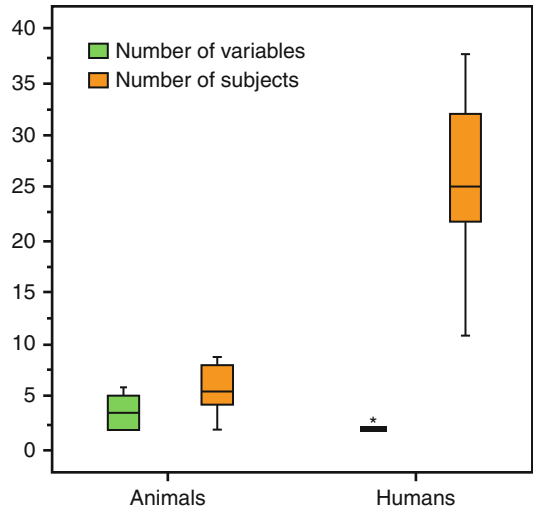


Fig. 21.79 Relationship between number of variables tested and number of subjects in animal experiments and human trials (Faggion et al. 2010. Reprinted with permission from Elsevier)

also be always considered when studies on efficacy and effectiveness are planned (Faggion et al. 2009).

References

- Aaboe M, Pinholt EM, Hjørting-Hansen E. Unicortical critical size defect of rabbit tibia is larger than 8 mm. *J Craniofac Surg.* 1994;5:201–3.
- Abukawa H, Zhang W, Young CS, Asrican R, Vacanti JP, Kaban LB, Troulis MJ, Yelick PC. Reconstructing mandibular defects using autologous tissue-engineered tooth and bone constructs. *J Oral Maxillofac Surg.* 2009;67:335–47.
- Achong R, Nishimura I, Ramachandran H, Howell TH, Fiorellini JP, Karimbux NY. Membrane type (MT) 1-matrix metalloproteinase (MMP) and MMP-2 expression in ligature-induced periodontitis in the rat. *J Periodontol.* 2003;74:494–500.
- Aizawa-Abe M, Ogawa Y, Masuzaki H, et al. Pathophysiological role of leptin in obesity-related hypertension. *J Clin Invest.* 2000;105:1243–52.
- Akizuki T, Oda S, Komaki M, Tsuchioka H, Kawakatsu N, Kikuchi A, Yamato M, Okano T, Ishikawa I. Application of periodontal ligament cell sheet for periodontal regeneration: a pilot study in beagle dogs. *J Periodontol Res.* 2005;40:245–51.
- Almeida JM, Theodoro LH, Bosco AF, Nagata MJ, Oshiiwa M, Garcia VG. In vivo effect of photodynamic therapy on periodontal bone loss in dental furcations. *J Periodontol.* 2008;79(6):1081–8.

- Alayan J, Ivanovski S, Gemmell E, Ford P, Hamlet S, Farah CS. Deficiency of iNOS contributes to *Porphyromonas gingivalis*-induced tissue damage. *Oral Microbiol Immunol*. 2006;21:360–5.
- Albuquerque C, Morinha F, Requicha J, Martins T, Dias I, Guedes-Pinto H, Bastos E, Viegas C. Canine periodontitis: the dog as an important model for periodontal studies. *Vet J*. 2012;191(3):299–305. Epub 2011 Sep 21.
- Alhezaime K, Al-Shalan T, O'Neill R, Shapurian T, Naghshbandi J, Levi Jr P, Griffin T. Connective tissue-cementum regeneration: a new histologic regeneration following the use of enamel matrix derivative in dehiscence-type defects. A dog model. *Int J Periodontics Restorative Dent*. 2009;29:425–33.
- Alikhani M, Alikhani Z, He H, Liu R, Popek BI, Graves DT. Lipopolysaccharides indirectly stimulate apoptosis and global induction of apoptotic genes in fibroblasts. *J Biol Chem*. 2003;278:52901–8.
- Al-Mashat HA, Kandru S, Liu R, Behl Y, Desta T, Graves DT. Diabetes enhances mRNA levels of proapoptotic genes and caspase activity, which contribute to impaired healing. *Diabetes*. 2006;55:487–95.
- Alnaeeli M, Penninger JM, Teng YT. Immune interactions with CD4+ T cells promote the development of functional osteoclasts from murine CD11c+ dendritic cells. *J Immunol*. 2006;177:3314–26.
- Alnaeeli M, Park J, Mahamed D, Penninger JM, Teng YT. Dendritic cells at the osteo-immune interface: implications for inflammation-induced bone loss. *J Bone Miner Res*. 2007;22:775–80.
- Al-Qareer AH, Afsah MR, Müller HP. A sheep cadaver model for demonstration and training periodontal surgical methods. *Eur J Dent Educ*. 2004;8:78–83.
- Arai K, Tanaka S, Yamamoto-Sawamura T, Sone K, Miyaishi O, Sumi Y. Aging changes in the periodontal bone of F344/N rat. *Arch Gerontol Geriatr*. 2005;40:225–9.
- Araujo MG, Lindhe J. GTR treatment of degree III furcation defects following application of enamel matrix proteins. An experimental study in dogs. *J Clin Periodontol*. 1998;25:524–30.
- Araújo MG, Lindhe J. Dimensional ridge alterations following tooth extraction. An experimental study in the dog. *J Clin Periodontol*. 2005;32:212–8.
- Araujo MG, Berglundh T, Lindhe J. GTR treatment of degree III furcation defects with 2 different resorbable barriers: an experimental study in dogs. *J Clin Periodontol*. 1998;25:253–9.
- Araujo MG, Berglundh T, Albrektsson T, Lindhe J. Bone formation in furcation defects. An experimental study in the dog. *J Clin Periodontol*. 1999;26:643–52.
- Araújo MG, Sukekava F, Wennström JL, Lindhe J. Ridge alterations following implant placement in fresh extraction sockets: an experimental study in the dog. *J Clin Periodontol*. 2005;32:645–52.
- Arce RM, Barros SP, Wacker B, Peters B, Moss K, Offenbacher S. Increased TLR4 expression in murine placentas after oral infection with periodontal pathogens. *Placenta*. 2009;30:156–62.
- Artzi Z, Weinreb M, Tal H, Nemcovsky CE, Rohrer MD, Prasad HS, Kozlovsky A. Experimental intrabony and periodontal defects treated with natural mineral combined with a synthetic cell-binding Peptide in the canine: morphometric evaluations. *J Periodontol*. 2006;77:1658–64.
- Baik DH, Hwang SJ, Kim CS, Lee YK, Cho CS, Chai JK, Kim CK, Choi SH. The effects of noncrystalline calcium phosphate glass on the healing of 1-wall intrabony defects in beagle dogs. *J Korean Acad Periodontol*. 2004;34:113–20.
- Bain JL, Lester SR, Henry WD, Bishop CM, Turnage AA, Naftel JP, Johnson RB. Comparative gender differences in local and systemic concentrations of proinflammatory cytokines in rats with experimental periodontitis. *J Periodontol Res*. 2009;44:133–40.
- Baker PJ. Genetic control of the immune response in pathogenesis. *J Periodontol*. 2005;76(11 Suppl):2042–6.
- Baker PJ, Evans RT, Roopenian DC. Oral infection with *Porphyromonas gingivalis* and induced alveolar bone loss in immunocompetent and severe combined immunodeficient mice. *Arch Oral Biol*. 1994;39:1035–40.
- Baker PJ, Carter S, Dixon M, Evans RT, Roopenian DC. Serum antibody response to oral infection precedes but does not prevent *Porphyromonas gingivalis* induced alveolar bone loss in mice. *Oral Microbiol Immunol*. 1999a;14:194–6.
- Baker PJ, Dixon M, Evans RT, Dufour L, Johnson E, Roopenian DC. CD4(+) T cells and the proinflammatory cytokines gamma interferon and interleukin-6 contribute to alveolar bone loss in mice. *Infect Immun*. 1999b;67:2804–9.
- Baker PJ, Dixon M, Evans RT, Roopenian DC. Heterogeneity of *Porphyromonas gingivalis* strains in the induction of alveolar bone loss in mice. *Oral Microbiol Immunol*. 2000a;15:27–32.
- Baker PJ, Dixon M, Roopenian DC. Genetic control of susceptibility to *Porphyromonas gingivalis*-induced alveolar bone loss in mice. *Infect Immun*. 2000b;68:5864–8.
- Baron R, Saffar JL. A quantitative study of bone remodeling during experimental periodontal disease in the golden hamster. *J Periodontol Res*. 1978;13:309–15.
- Baroukh B, Saffar JS. The effect of leukotriene synthesis inhibitors on hamster periodontitis. *Arch Oral Biol*. 1990;35(Suppl):189S–92.
- Baroukh B, Saffar JL. Identification of osteoclasts and their mononuclear precursors. A comparative histological and histochemical study in hamster periodontitis. *J Periodontol Res*. 1991;26(3 Pt 1):161–6.
- Barros SP, Arce RM, Galloway P, Lawter JR, Offenbacher S. Therapeutic effect of a topical CCR2 antagonist on induced alveolar bone loss in mice. *J Periodontol Res*. 2011;46:246–51.
- Bartold PM, Marshall RI, Haynes DR. Periodontitis and rheumatoid arthritis: a review. *J Periodontol*. 2005;76(11 Suppl):2066–74.
- Bartold PM, Marino V, Cantley M, Haynes DR. Effect of *Porphyromonas gingivalis*-induced inflammation on

- the development of rheumatoid arthritis. *J Clin Periodontol.* 2010;37:405–11.
- Behbehani MJ, Jordan HV, Heeley JD. Oral colonization and pathogenicity of *Actinomyces israelii* in gnotobiotic rats. *J Dent Res.* 1983;62:69–74.
- Benatti BB, César-Neto JB, Gonçalves PF, Sallum EA, Nociti Jr FH. Smoking affects the self-healing capacity of periodontal tissues. A histological study in the rat. *Eur J Oral Sci.* 2005;113:400–3.
- Benatti BB, César-Neto JB, Casati MZ, Sallum EA, Sallum AW, Nociti Jr FH. Periodontal healing may be affected by aging: a histologic study in rats. *J Periodontol Res.* 2006;41:329–33.
- Bendyk A, Marino V, Zilm PS, Howe P, Bartold PM. Effect of dietary omega-3 polyunsaturated fatty acids on experimental periodontitis in the mouse. *J Periodontol Res.* 2009;44:211–6.
- Bergenholtz G, Wikesjö UM, Sorensen RG, Xiropaidis AV, Wozney JM. Observations on healing following endodontic surgery in nonhuman primates (*Macaca fascicularis*): effects of rhBMP-2. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2006;101:116–25.
- Björnsson MJ, Velschow S, Stoltze K, Havemose-Poulsen A, Schou S, Holmstrup P. The influence of diet consistency, drinking water and bedding on periodontal disease in Sprague-Dawley rats. *J Periodontol Res.* 2003;38:543–50.
- Björnsson MJ, Havemose-Poulsen A, Stoltze K, Holmstrup P. Influence of the matrix metalloproteinase inhibitor batimastat (BB-94) on periodontal bone destruction in Sprague-Dawley rats. *J Periodontol Res.* 2004;39:269–74.
- Blumenthal NM, Alves ME, Al-Huwais S, Hofbauer AM, Koperski RD. Defect-determined regenerative options for treating periodontal intrabony defects in baboons. *J Periodontol.* 2003;74:10–24.
- Bobetsis YA, Barros SP, Offenbacher S. Exploring the relationship between periodontal disease and pregnancy complications. *J Am Dent Assoc.* 2006;137zc(Suppl):7S–13.
- Boyce BF, Aufdemorte TB, Garrett IR, Yates AJ, Mundy GR. Effects of interleukin-1 on bone turnover in normal mice. *Endocrinology.* 1989;125:1142–50.
- Bezerra MM, de Lima V, Alencar VB, Vieira IB, Brito GA, Ribeiro RA, Rocha FA. Selective cyclooxygenase-2 inhibition prevents alveolar bone loss in experimental periodontitis in rats. *J Periodontol.* 2000 Jun;71(6):1009–14.
- Branch-Mays GL, Dawson DR, Gunsolley JC, Reynolds MA, Ebersole JL, Novak KF, Mattison JA, Ingram DK, Novak MJ. The effects of a calorie-reduced diet on periodontal inflammation and disease in a non-human primate model. *J Periodontol.* 2008;79:1184–91.
- Brecx M, Patters MR. Morphology of polymorphonuclear neutrophils during periodontal disease in the cynomolgus monkey. *J Clin Periodontol.* 1985;12:591–606.
- Brecx MC, Nalbandian J, Ooya K, Kornman KS, Robertson PB. Morphological studies on periodontal disease in the cynomolgus monkey. II. Light microscopic observations on ligature-induced periodontitis. *J Periodontol Res.* 1985;20:165–75.
- Breivik T, Opstad PK, Gjermo P, Thrane PS. Effects of hypothalamic-pituitary-adrenal axis reactivity on periodontal tissue destruction in rats. *Eur J Oral Sci.* 2000a;108:115–22.
- Breivik T, Sluyter F, Hof M, Cools A. Differential susceptibility to periodontitis in genetically selected Wistar rat lines that differ in their behavioral and endocrinological response to stressors. *Behav Genet.* 2000b;30:123–30.
- Breivik T, Gundersen Y, Myhrer T, Fonnum F, Osmundsen H, Murison R, et al. Enhanced susceptibility to periodontitis in an animal model of depression: reversed by chronic treatment with the anti-depressant tianeptine. *J Clin Periodontol.* 2006;33:469–77.
- Brodala N, Merricks EP, Bellinger DA, Damrongsri D, Offenbacher S, Beck J, et al. *Porphyromonas gingivalis* bacteremia induces coronary and aortic atherosclerosis in normocholesterolemic and hypercholesterolemic pigs. *Arterioscler Thromb Vasc Biol.* 2005;25:1446–51.
- Burckhardt JJ, Gaegauf-Zollinger R, Schmid R, Guggenheim B. Alveolar bone loss in rats after immunization with *Actinomyces viscosus*. *Infect Immun.* 1981;31:971–7.
- Burns E, Bachrach G, Shapira L, Nussbaum G. Cutting Edge: TLR2 is required for the innate response to *Porphyromonas gingivalis*: activation leads to bacterial persistence and TLR2 deficiency attenuates induced alveolar bone resorption. *J Immunol.* 2006;177(12):8296–300.
- Busch L, Sterin-Borda L, Borda E. Beta-adrenoceptor alterations coupled with secretory response and experimental periodontitis in rat submandibular glands. *Arch Oral Biol.* 2008;16:53–509.
- Buser D, Hoffmann B, Bernard JP, Lussi A, Mettler D, Schenk RK. Evaluation of filling materials in membrane – protected bone defects. A comparative histomorphometric study in the mandible of miniature pigs. *Clin Oral Implants Res.* 1998;9:137–50.
- Caffesse RG, Dominguez LE, Nasjleti CE, Castelli WA, Morrison EC, Smith BA. Furcation defects in dogs treated by guided tissue regeneration (GTR). *J Periodontol.* 1990;61:45–50.
- Calvo-Guirado JL, Delgado-Ruiz RA, Ramírez-Fernández MP, Maté-Sánchez JE, Ortiz-Ruiz A, Marcus A. Histomorphometric and mineral degradation study of Ossceram: a novel biphasic B-tricalcium phosphate, in critical size defects in rabbits. *Clin Oral Implants Res.* 2012;23:667–75.
- Cantley MD, Bartold PM, Marino V, Reid RC, Fairlie DP, Wyszynski RN, Zilm PS, Haynes DR. The use of live-animal micro-computed tomography to determine the effect of a novel phospholipase A2 inhibitor on alveolar bone loss in an in vivo mouse model of periodontitis. *J Periodontol Res.* 2009;44:317–22.
- Cantley MD, Haynes DR, Marino V, Bartold PM. Pre-existing periodontitis exacerbates experimental arthritis

- in a mouse model. *J Clin Periodontol.* 2011;38:532–41.
- Caplanis N, Lee MB, Zimmerman GJ, Selvig KA, Wikesjö UM. Effect of allogenic freeze-dried demineralized bone matrix on guided tissue regeneration in dogs. *J Periodontol.* 1998;69:851–6.
- Cappelli D, Steffen MJ, Holt SC, Ebersole JL. Periodontitis in pregnancy: clinical and serum antibody observations from a baboon model of ligature-induced disease. *J Periodontol.* 2009;80:1154–65.
- Casati MZ, Sallum EA, Caffesse RG, Nociti Jr FH, Sallum AW, Pereira SL. Guided tissue regeneration with a bioabsorbable polylactic acid membrane in gingival recessions: a histometric study in dogs. *J Periodontol.* 2000;71:238–48.
- Caton J, Mota L, Gandini L, Laskaris B. Non-human primate models for testing the efficacy and safety of periodontal regeneration procedures. *J Periodontol.* 1994;65:1143–50.
- César-Neto JB, Benatti BB, Sallum EA, Casati MZ, Nociti Jr FH. The influence of cigarette smoke inhalation and its cessation on the tooth-supporting alveolar bone: a histometric study in rats. *J Periodontal Res.* 2006;41:118–23.
- Chang KM, Ramamurthy NS, McNamara TF, Genco RJ, Golub LM. Infection with a gram-negative organism stimulates gingival collagenase production in non-diabetic and diabetic germfree rats. *J Periodontal Res.* 1988;23:239–44.
- Chang KM, Ramamurthy NS, McNamara TF, Evans RT, Klausen B, Murray PA, Golub LM. Tetracyclines inhibit *Porphyromonas gingivalis*-induced alveolar bone loss in rats by a non-antimicrobial mechanism. *J Periodontal Res.* 1994;29:242–9.
- Cheng W-C, Huang R-Y, Chiang C-Y, Chen J-K, Liu C-H, Chu C-L, Fu E. Ameliorative effect of quercetin on the destruction caused by experimental periodontitis in rats. *J Periodontal Res.* 2010;45:788–95.
- Chiang CY, Kyritsis G, Graves DT, Amar S. Interleukin-1 and tumor necrosis factor activities partially account for calvarial bone resorption induced by local injection of lipopolysaccharide. *Infect Immun.* 1999;67:4231–6.
- Choi SY, Nilvéus RE, Minutello RD, Zimmerman GJ, Wikesjö UM. Effect of a collagen matrix on healing in periodontal fenestration defects in dogs. *J Periodontol.* 1993;64:878–82.
- Choi SH, Kim CK, Cho KS, Huh JS, Sorensen RG, Wozney JM, Wikesjö UM. Effect of recombinant human bone morphogenetic protein-2/absorbable collagen sponge (rhBMP-2/ACS) on healing in 3-wall intrabony defects in dogs. *J Periodontol.* 2002;73:63–72.
- Collins JG, Smith MA, Arnold RR, Offenbacher S. Effects of *Escherichia coli* and *Porphyromonas gingivalis* lipopolysaccharide on pregnancy outcome in the golden hamster. *Infect Immun.* 1994a;62:4652–5.
- Collins JG, Windley 3rd HW, Arnold RR, Offenbacher S. Effects of a *Porphyromonas gingivalis* infection on inflammatory mediator response and pregnancy outcome in hamsters. *Infect Immun.* 1994b;62:4356–61.
- Corrêa MG, Campos ML, Benatti BB, Marques MR, Casati MZ, Nociti Jr FH, Sallum EA. The impact of cigarette smoke inhalation on the outcome of enamel matrix derivative treatment in rats: histometric analysis. *J Periodontol.* 2010;81:1820–8.
- Craig RG, Kallur SP, Inoue M, Rosenberg PA, LeGeros RZ. Effect of enamel matrix proteins on the periodontal connective tissue-material interface after wound healing. *J Biomed Mater Res A.* 2004;69:180–7.
- Craig RG, Kamer AR, Kallur SP, Inoue M, Tarnow DP. Effects of periodontal cell grafts and enamel matrix proteins on the implant connective tissue interface: a pilot study in the minipig. *J Oral Implantol.* 2006;32:228–36.
- Crawford JM, Taubman MA, Smith DJ. The natural history of periodontal bone loss in germfree and gnotobiotic rats infected with periodontopathic microorganisms. *J Periodontal Res.* 1978;13:316–25.
- da Monteiro Silva AM, Newman HN, Oakley DA. Psychosocial factors in inflammatory periodontal diseases. A review. *J Clin Periodontol.* 1995;22:516–26.
- da Silva Pereira SL, Sallum AW, Casati MZ. Comparison of bioabsorbable and non-resorbable membranes in the treatment of dehiscence-type defects: a histomorphometric study in dogs. *J Periodontol.* 2000;71:1306–14.
- Danesh-Meyer MJ, Pack AR, McMillan MD. A comparison of 2 polytetrafluoroethylene membranes in guided tissue regeneration in sheep. *J Periodontal Res.* 1997;32(1 Pt 1):20–30.
- de Aquino SG, Guimaraes MR, Stach-Machado DR, da Silva JA, Spolidorio LC, Rossa Jr C. Differential regulation of MMP-13 expression in two models of experimentally induced periodontal disease in rats. *Arch Oral Biol.* 2009;54:609–17.
- DeCarlo AA, Huang Y, Collyer CA, Langley DB, Katz J. Feasibility of an HA2 domain-based periodontitis vaccine. *Infect Immun.* 2003;71:562–6.
- Decker JF, Lee J, Cortella CA, Polimeni G, Rohrer MD, Wozney JM, Hall J, Susin C, Wikesjö UM. Evaluation of implants coated with recombinant human bone morphogenetic protein-2 and vacuum-dried using the critical-size supraalveolar peri-implant defect model in dogs. *J Periodontol.* 2010;81:1839–49.
- DeLaurier A, Boyde A, Jackson B, Horton MA, Price JS. Identifying early osteoclastic resorptive lesions in feline teeth: a model for understanding the origin of multiple idiopathic root resorption. *J Periodontal Res.* 2009;44:248–57.
- Di Paola R, Mazzon E, Muia' C, Terrana D, Greco S, Britti D, Santori D, Oteri G, Cordasco G, Cuzzocrea S. 5-aminoisoquinolin-1(2H)-one, a water soluble poly (ADP-ribose) polymerase (PARP) inhibitor reduces the evolution of experimental periodontitis in rats. *J Clin Periodontol.* 2007;34:95–102.
- Ding G, Liu Y, Wang W, Wei F, Liu D, Fan Z, An Y, Zhang C, Wang S. Allogeneic periodontal ligament stem cell

- therapy for periodontitis in swine. *Stem Cells*. 2010;28:1829–38.
- Dogan A, Ozdemir A, Kubar A, Oygür T. Assessment of periodontal healing by seeding of fibroblast-like cells derived from regenerated periodontal ligament in artificial furcation defects in a dog: a pilot study. *Tissue Eng*. 2002;8:273–82.
- Dogan A, Ozdemir A, Kubar A, Oygur T. Healing of artificial fenestration defects by seeding of fibroblast-like cells derived from regenerated periodontal ligament in a dog: a preliminary study. *Tissue Eng*. 2003;9:1189–96.
- Donos N, Sculean A, Glavind L, Reich E, Karring T. Wound healing of degree III furcation involvements following guided tissue regeneration and/or Emdogain. A histologic study. *J Clin Periodontol*. 2003;30:1061–8.
- Doxey DL, Cutler CW, Iacopino AM. Diabetes prevents periodontitis-induced increases in gingival platelet derived growth factor-B and interleukin 1-beta in a rat model. *J Periodontol*. 1998;69:113–9.
- Dreyer WP, Basson NJ. *Porphyromonas gingivalis* as putative pathogen in ovine periodontitis. *J Dent Assoc S Afr*. 1992;47:513–6.
- Duarte PM, Tezolin KR, Figueiredo LC, Feres M, Bastos MF. Microbial profile of ligature-induced periodontitis in rats. *Arch Oral Biol*. 2010;55:142–7.
- Dumitrescu AL, Abd-El-Aleem S, Morales-Aza B, Donaldson LF. A model of periodontitis in the rat: effect of lipopolysaccharide on bone resorption, osteoclast activity, and local peptidergic innervation. *J Clin Periodontol*. 2004;31:596–603.
- Duncan WJ, Persson GR, Sims TJ, Braham P, Pack ARC, Page RC. Ovine periodontitis as a potential model for periodontal studies. Cross-sectional analysis of clinical, microbiological, and serum immunological parameters. *J Clin Periodontol*. 2003;30:63–72.
- Dyer BL, Caffesse RG, Nasjleti CE, Morrison EC. Guided tissue regeneration with dentin biomodification. *J Periodontol*. 1993;64:1052–60.
- Ebersole JL, Kraig E, Bauman G, Spitznagel JK, Kolodrubetz D. Molecular approaches to leucotoxin as a virulence component in *Actinobacillus actinomycetemcomitans*. *Arch Oral Biol*. 1990;35(Suppl):69S–78.
- Ebersole JL, Brunsvold M, Steffensen B, Wood R, Holt SC. Effects of immunization with *Porphyromonas gingivalis* and *Prevotella intermedia* on progression of ligature-induced periodontitis in the nonhuman primate *Macaca fascicularis*. *Infect Immun*. 1991;59:3351–9.
- Ebersole JL, Cappelli D, Holt SC, Singer RE, Filloon T. Gingival crevicular fluid inflammatory mediators and bacteriology of gingivitis in nonhuman primates related to susceptibility to periodontitis. *Oral Microbiol Immunol*. 2000;15:19–26.
- Ebersole JL, Cappelli D, Mathys EC, Steffen MJ, Singer RE, Montgomery M, Mott GE, Novak MJ. Periodontitis in humans and non-human primates: oral-systemic linkage inducing acute phase proteins. *Ann Periodontol*. 2002;7:102–11.
- Ebersole JL, Steffen MJ, Holt SC, Kesavalu L, Chu L, Cappelli D. Systemic inflammatory responses in progressing periodontitis during pregnancy in a baboon model. *Clin Exp Immunol*. 2010;162:550–9.
- Ekuni D, Firth JD, Putnins EE. Regulation of epithelial cell growth factor receptor protein and gene expression using a rat periodontitis model. *J Periodontol Res*. 2006;41:340–9.
- Ekuni D, Tomofuji T, Tamaki N, Sanbe T, Azuma T, Yamanaka R, Yamamoto T, Watanabe T. Mechanical stimulation of gingiva reduces plasma 8-OHdG level in rat periodontitis. *Arch Oral Biol*. 2008;53:324–9.
- Ekuni D, Tomofuji T, Sanbe T, Irie K, Azuma T, Maruyama T, Tamaki N, Murakami J, Koikeguchi S, Yamamoto T. Periodontitis-induced lipid peroxidation in rat descending aorta is involved in the initiation of atherosclerosis. *J Periodontol Res*. 2009;44:434–42.
- Ekuni D, Tomofuji T, Irie K, Kasuyama K, Umakoshi M, Azuma T, Tamaki N, Sanbe T, Endo Y, Yamamoto T, Nishida T, Morita M. Effects of periodontitis on aortic insulin resistance in an obese rat model. *Lab Invest*. 2010;90:348–59.
- Elharar F, Rodriguez HJ, Benque EP, Caffesse RG. Guided tissue regeneration with bioabsorbable and expanded polytetrafluoroethylene barrier membranes in the treatment of naturally occurring periodontal defects in dogs. *J Periodontol*. 1998;69:1218–28.
- Endo Y, Tomofuji T, Ekuni D, Irie K, Azuma T, Tamaki N, Yamamoto T, Morita M. Experimental periodontitis induces gene expression of proinflammatory cytokines in liver and white adipose tissues in obesity. *J Periodontol*. 2010;81:520–6.
- Escartin Q, Lallam-Laroye C, Baroukh B, Morvan FO, Caruelle JP, Godeau G, Barritault D, Saffar JL. A new approach to treat tissue destruction in periodontitis with chemically modified dextran polymers. *FASEB J*. 2003;17:644–51.
- Evans RT, Klausen B, Sojar HT, Bedi GS, Sfintescu C, Ramamurthy NS, Golub LM, Genco RJ. Immunization with *Porphyromonas* (Bacteroides) *gingivalis* fimbriae protects against periodontal destruction. *Infect Immun*. 1992;60:2926–35.
- Faggion CM Jr, Listl S, Giannakopoulos NN. The methodological quality of systematic reviews of animal studies in dentistry. *Vet J*. 2012;192:140–7.
- Faggion Jr CM, Schmitter M, Tu YK. Assessment of replication of research evidence from animals to humans in studies on peri-implantitis therapy. *J Dent*. 2009;37:737–47.
- Faggion Jr CM, Chambrone L, Gondim V, Schmitter M, Tu YK. Comparison of the effects of treatment of peri-implant infection in animal and human studies: systematic review and meta-analysis. *Clin Oral Implants Res*. 2010;21:137–47.
- Faggion Jr CM, Giannakopoulos NN, Listl S. Risk of bias of animal studies on regenerative procedures for periodontal and peri-implant bone defects – a systematic review. *J Clin Periodontol*. 2011;38:1154–60.
- Felipe ME, Andrade PF, Novaes Jr AB, Grisi MF, Souza SL, Taba Jr M, Palioto DB. Potential of bioactive glass

- particles of different size ranges to affect bone formation in interproximal periodontal defects in dogs. *J Periodontol.* 2009;80:808–15.
- Fernandes MI, Gaio EJ, Oppermann RV, Rados PV, Rosing CK. Comparison of histometric and morphometric analyses of bone height in ligature-induced periodontitis in rats. *Braz Oral Res.* 2007;21:216–21.
- Fernandes LA, Martins TM, de Almeida JM, Nagata MJ, Theodoro LH, Garcia VG, Bosco AF. Experimental periodontal disease treatment by subgingival irrigation with tetracycline hydrochloride in rats. *J Appl Oral Sci.* 2010;18:635–40.
- Fischer RG, Klinge B. Clinical and histological evaluation of ligature-induced periodontitis in the domestic ferret. *J Clin Periodontol.* 1994;21:230–9.
- Fischer RG, Edwardsson S, Klinge B. Oral microflora of the ferret at the gingival sulcus and mucosa membrane in relation to ligature-induced periodontitis. *Oral Microbiol Immunol.* 1994;9:40–9.
- Fisher RG, Klinge B, Atstrom R. Normal histologic features of domestic ferret periodontium. *Scand J Dent Res.* 1993;101:357–62.
- Friskens KW, Laws AJ, Tagg JR, Orr MB. Environmental influences on the progression of clinical and microbiological parameters of sheep periodontal disease. *Res Vet Sci.* 1989;46:147–52.
- Fujishiro N, Anan H, Hamachi T, Maeda K. The role of macrophages in the periodontal regeneration using Emdogain® gel. *J Periodontal Res.* 2008;43:143–55.
- Garant PR. An electron microscopic study of the periodontal tissues of germfree rats and rats monoinfected with *Actinomyces naeslundii*. *J Periodontal Res.* 1976a;15:3–79.
- Garant PR. Light and electron microscopic observations of osteoclastic alveolar bone resorption in rats monoinfected with *Actinomyces naeslundii*. *J Periodontol.* 1976b;47:717–23.
- Garlet GP, Cardoso CR, Silva TA, Ferreira BR, Avila-Campos MJ, Cunha FQ, Silva JS. Cytokine pattern determines the progression of experimental periodontal disease induced by *Actinobacillus actinomycetemcomitans* through the modulation of MMPs, RANKL, and their physiological inhibitors. *Oral Microbiol Immunol.* 2006;21:12–20.
- Garlet GP, Cardoso CR, Mariano FS, Claudino M, de Assis GF, Campanelli AP, Ávila-Campos MJ, Silva JS. Regulatory T cells attenuate experimental periodontitis progression in mice. *J Clin Periodontol.* 2010;37:591–600.
- Gašperšič R, Štibilar-Martinčič D, Skalerič U. Influence of restraint stress on ligature-induced periodontitis in rats. *Eur J Oral Sci.* 2002;110:125–9.
- Gavin JB. Ultrastructural features of chronic marginal gingivitis. *J Periodontal Res.* 1970;5:19–29.
- Gemmell E, Bird PS, Ford PJ, Ashman RB, Gosling P, Hu Y, Seymour GJ. Modulation of the antibody response by *Porphyromonas gingivalis* and *Fusobacterium nucleatum* in a mouse model. *Oral Microbiol Immunol.* 2004;19:247–51.
- Genco CA, Cutler CW, Kapczynski D, Maloney K, Arnold RR. A novel mouse model to study the virulence of and host response to *Porphyromonas* (Bacteroides) *gingivalis*. *Infect Immun.* 1991;59:1255–63.
- Genco CA, Kapczynski DR, Cutler CW, Arko RJ, Arnold RR. Influence of immunization on *Porphyromonas gingivalis* colonization and invasion in the mouse chamber model. *Infect Immun.* 1992;60:1447–54.
- Genco CA, Odusanya BM, Potempa J, Mikolajczyk-Pawlinska J, Travis J. A peptide domain on gingipain R which confers immunity against *Porphyromonas gingivalis* infection in mice. *Infect Immun.* 1998a;66:4108–14.
- Genco CA, Van Dyke T, Amar S. Animal models for *Porphyromonas gingivalis*-mediated periodontal disease. *Trends Microbiol.* 1998b;6:444–9.
- Germanier Y, Tosatti S, Broggin N, Textor M, Buser D. Enhanced bone apposition around biofunctionalized sandblasted and acid etched titanium implant surfaces. A histomorphometric study in miniature pigs. *Clin Oral Implants Res.* 2006;17:251–7.
- Giannobile WV, Finkelman RD, Lynch SE. Comparison of canine and non-human primate animal models for periodontal regenerative therapy: Results following a single administration of PDGF/IGF-I. *J Periodontol.* 1994;65:1158–68.
- Gibson 3rd FC, Genco CA. Prevention of *Porphyromonas gingivalis*-induced oral bone loss following immunization with gingipain R1. *Infect Immun.* 2001;69:7959–63.
- Gibson 3rd FC, Hong C, Chou HH, Yumoto H, Chen J, Lien E, Wong J, Genco CA. Innate immune recognition of invasive bacteria accelerates atherosclerosis in apolipoprotein E-deficient mice. *Circulation.* 2004;109:2801–6.
- Gineste L, Gineste M, Bluche L, et al. Histomorphometric comparison of three bioabsorbable GTR barrier membranes in the canine model. *Int J Periodontics Restorative Dent.* 2005a;25:61–71.
- Gineste L, Gineste M, Bluche L, Guilhem A, Elefterion A, Frayssinet P, Durand D, Wang HL. Histomorphometric comparison of three bioabsorbable GTR barrier membranes in the canine model. *Int J Periodontics Restorative Dent.* 2005b;25:61–71.
- Gitlin JM, Loftin CD. Cyclooxygenase-2 inhibition increases lipopolysaccharide-induced atherosclerosis in mice. *Cardiovasc Res.* 2009;81:400–7.
- Garant PR, Cho MI. Histopathogenesis of spontaneous periodontal disease in conventional rats. II. Ultrastructural features of the inflamed subepithelial connective tissue. *J Periodontal Res.* 1979a;14(4):310–22.
- Garant PR, Cho MI. Histopathogenesis of spontaneous periodontal disease in conventional rats. I. Histometric and histologic study. *J Periodontal Res.* 1979b;14(4):297–309.
- Graves DT, Oskoui M, Volejnikova S, Naguib G, Cai S, Desta T, Kakouras A, Jiang Y. Tumor necrosis factor modulates fibroblast apoptosis, PMN recruitment, and osteoclast formation in response to *P. gingivalis* infection. *J Dent Res.* 2001;80:1875–9.
- Graves DT, Naguib G, Lu H, Desta T, Amar S. *Porphyromonas gingivalis* fimbriae are pro-

- inflammatory but do not play a prominent role in the innate immune response to *P. gingivalis*. *J Endotoxin Res*. 2005a;11:13–8.
- Graves DT, Naguib G, Lu H, Leone C, Hsue H, Krall E. Inflammation is more persistent in type 1 diabetic mice. *J Dent Res*. 2005b;84:324–8.
- Graves DT, Fine D, Teng Y-T, Van Dyke TE, Hajishengallis G. The use of rodent models to investigate host-bacteria interactions related to periodontal diseases. *J Clin Periodontol*. 2008;35:89–105.
- Graves DT, Kang J, Andriankaja O, Wada K, Rossa Jr C. Animal models to study host-bacteria interactions involved in periodontitis. *Front Oral Biol*. 2012;15:117–32.
- Graziani F, Laurell L, Tonetti M, Gottlow J, Berglundh T. Periodontal wound healing following GTR therapy of dehiscence-type defects in the monkey: short-, medium- and long-term healing. *J Clin Periodontol*. 2005;32:905–14.
- Guiha R, el Khodairy S, Mota L, Caffesse R. Histological evaluation of healing and revascularization of the sub-epithelial connective tissue graft. *J Periodontol*. 2001;72:470–8.
- Gaegauf-Zollinger R, Burckhardt JJ, Guggenheim B. Radiographic measurements of alveolar bone loss in the rat. *Arch Oral Biol*. 1982;27(8):651–8.
- Hachicha M, Pouliot M, Petasis NA, Serhan CN. Lipoxin (LX)A4 and aspirin-triggered 15-epi-LXA4 inhibit tumor necrosis factor 1 α -initiated neutrophil responses and trafficking: regulators of a cytokine-chemokine axis. *J Exp Med*. 1999;189:1923–30.
- Hackam DG, Redelmeier DA. Translation of research evidence from animals to humans. *JAMA*. 2006;296:1731–2.
- Hajishengallis G, Shakhathreh MA, Wang M, Liang S. Complement receptor 3 blockade promotes IL-12-mediated clearance of *Porphyromonas gingivalis* and negates its virulence in vivo. *J Immunol*. 2007;179:2359–67.
- Hayashi C, Gudino CV, Gibson FC 3rd, Genco CA. Review: Pathogen-induced inflammation at sites distant from oral infection: bacterial persistence and induction of cell-specific innate immune inflammatory pathways. *Mol Oral Microbiol*. 2010;25(5):305–16.
- Hammarström L, Heijl L, Gestrelus S. Periodontal regeneration in a buccal dehiscence model in monkeys after application of enamel matrix proteins. *J Clin Periodontol*. 1997;24:669–77.
- Haney JM, Zimmerman GJ, Wikesjö UME. Periodontal repair in dogs: evaluation of the natural disease model. *J Clin Periodontol*. 1995;22:208–13.
- Hanisch O, Sorensen RG, Kinoshita A, Spiekermann H, Wozney JM, Wikesjö UM. Effect of recombinant human bone morphogenetic protein-2 in dehiscence defects with non-submerged immediate implants: an experimental study in Cynomolgus monkeys. *J Periodontol*. 2003;74:648–57.
- Hardham J, Reed M, Wong J, King K, Laurinat B, Sfintescu C, Evans RT. Evaluation of a monovalent companion animal periodontal disease vaccine in an experimental mouse periodontitis model. *Vaccine*. 2005;23:3148–56.
- Harper DS, Mann PH, Regnier S. Measurement of dietary and dentifrice effects upon calculus accumulation rates in the domestic ferret. *J Dent Res*. 1990;69:447–50.
- Hardt AB. Bisphosphonate effects on alveolar bone during rat molar drifting. *J Dent Res*. 1988;67:1430–3.
- Hart GT, Shaffer DJ, Akilesh S, Brown AC, Moran L, Roopenian DC, Baker PJ. Quantitative gene expression profiling implicates genes for susceptibility and resistance to alveolar bone loss. *Infect Immun*. 2004;72:4471–9.
- Hasturk H, Kantarci A, Ebrahimi N, Andry C, Holick M, Jones VL, Van Dyke TE. Topical H2 antagonist prevents periodontitis in a rabbit model. *Infect Immun*. 2006a;74:2402–14.
- Hasturk H, Kantarci A, Ohira T, Arita M, Ebrahimi N, Chiang N, Petasis NA, Levy BD, Serhan CN, Van Dyke TE. RvE1 protects from local inflammation and osteoclast-mediated bone destruction in periodontitis. *FASEB J*. 2006b;20:401–3.
- Hasturk H, Jones VL, Andry C, Kantarci A. 1-Tetradecanol complex reduces progression of *Porphyromonas gingivalis*-induced experimental periodontitis in rabbits. *J Periodontol*. 2007a;78:924–32.
- Hasturk H, Kantarci A, Goguet-Surmenian E, Blackwood A, Andry C, Serhan CN, Van Dyke TE. Resolvin E1 regulates inflammation at the cellular and tissue level and restores tissue homeostasis in vivo. *J Immunol*. 2007b;179:7021–9.
- Hasturk H, Goguet-Surmenian E, Blackwood A, Andry C, Kantarci A. 1-Tetradecanol complex: therapeutic actions in experimental periodontitis. *J Periodontol*. 2009;80:1103–13.
- Hasturk H, Kantarci A, Ghattas M, Schmidt M, Giordano RA, Ashman A, Diekwisch TG, Van Dyke T. The use of light/chemically-hardened poly-methyl-methacrylate, poly-hydroxyl-ethyl-methacrylate and calcium hydroxide graft material in combination with polyanhydride around implants in minipigs: part I: immediate stability and function. *J Periodontol*. 2011;82:1339–52.
- Hasturk H, Kantarci A, Van Dyke TE. Paradigm shift in the pharmacological management of periodontal diseases. *Front Oral Biol*. 2012;15:160–76.
- Hautamäki MP, Aho AJ, Alander P, Rekola J, Gunn J, Strandberg N, Vallittu PK. Repair of bone segment defects with surface porous fiber-reinforced polymethyl methacrylate (PMMA) composite prosthesis: histomorphometric incorporation model and characterization by SEM. *Acta Orthop*. 2008;79:555–64.
- Hayashi C, Kinoshita A, Oda S, Mizutani K, Shirakata Y. Ishikawa Injectable calcium phosphate bone cement provides favorable space and a scaffold for periodontal regeneration in dogs. *J Periodontol*. 2006;77:940–6.
- Heijl L, Wennstrom J, Lindhe J, Socransky SS. Periodontal disease in gnatobiotic rats. *J Periodontal Res*. 1980;15:405–19.

- Hernández M, Gamonal J, Salo T, Tervahartiala T, Hukkanen M, Tjäderhane L, Sorsa T. Reduced expression of LIX/CXCL5 in *Porphyromonas gingivalis*-induced experimental periodontitis in matrix metalloproteinase-8 null mice. *J Periodontol Res*. 2011;46:58–66.
- Hickey JS, O'Neal RB, Scheidt MJ, Strong SL, Turgeon D, Van Dyke TE. Microbiologic characterization of ligature-induced peri-implantitis in the microswine model. *J Periodontol*. 1991;62:548–53.
- Hojó K, Tamura A, Mizoguchi C, Kato D, Ohshima T, Maeda N. Predominant bacteria recovered from a periodontitis site in a hamster model raised by silk-ligature with *Porphyromonas gingivalis* infection. *Biosci Biotechnol Biochem*. 2008;72:1348–51.
- Holt SC, Ebersole J, Felton J, Brunsvold M, Kornman KS. Implantation of *Bacteroides gingivalis* in nonhuman primates initiates progression of periodontitis. *Science*. 1988;239(4835):55–7.
- Holzhausen M, Rossa Júnior C, Marcantonio Júnior E, Nassar PO, Spolidório DM, Spolidório LC. Effect of selective cyclooxygenase-2 inhibition on the development of ligature-induced periodontitis in rats. *J Periodontol*. 2002;73:1030–6.
- Holzhausen M, Garcia DF, Pepato MT, Marcantonio Jr E. The influence of short-term diabetes mellitus and insulin therapy on alveolar bone loss in rats. *J Periodontol Res*. 2004;39:188–93.
- Hu SW, Huang CH, Huang HC, Lai YY, Lin YY. Transvascular dissemination of *Porphyromonas gingivalis* from a sequestered site is dependent upon activation of the kallikrein/kinin pathway. *J Periodontol Res*. 2006;41:200–7.
- Huang KK, Shen C, Chiang CY, Hsieh YD, Fu E. Effects of bone morphogenetic protein-6 on periodontal wound healing in a fenestration defect of rats. *J Periodontol Res*. 2005;40:1–10.
- Huang S, Lu F, Zhang Z, Yang X, Chen Y. The role of psychological stress induced hypoxia-inducible factor-1 α in rat experimental periodontitis. *J Periodontol*. 2010;82:934–41.
- Hyvärinen K, Tuomainen AM, Laitinen S, Bykov IL, Törmäkangas L, Lindros K, Käkälä R, Alfthan G, Salminen I, Jauhiainen M, Kovanen PT, Leinonen M, Saikku P, Pussinen PJ. Chlamydial and periodontal pathogens induce hepatic inflammation and fatty acid imbalance in apolipoprotein E-deficient mice. *Infect Immun*. 2009;77:3442–9.
- Ijuhin N, Miyauchi M, Ito H, Takata T, Ogawa I, Nikai H. Enhanced collagen phagocytosis by rat molar periodontal fibroblasts after topical application of lipopolysaccharide – ultrastructural observations and morphometric analysis. *J Periodontol Res*. 1992;27:167–75.
- Irokawa T, Ishida E, Fukuoka Y, Maita E, Yanagawa S. A study on the effects of continuous mechanical irritation to the periodontal tissues of germ-free and conventional rats (in Japanese). *Nippon Shishubyo Gakkai Kaishi*. 1977;19:333–40.
- Irving JT, Socransky SS, Heeley JD. Histological changes in experimental periodontal disease in gnotobiotic rats and conventional hamsters. *J Periodontol Res*. 1974;9:73–80.
- Irving JT, Newman MG, Socransky SS, Heeley JD. Histological changes in experimental periodontal disease in rats mono-infected with a gram-negative organism. *Arch Oral Biol*. 1975a;20:219–20.
- Irving JT, Heeley JD, Socransky SS. Cellular response to subgingival injection of bacterial products in the rat. *J Periodontol Res*. 1975b;10:324–31.
- Irving JT, Socransky SS, Tanner AC. Histological changes in experimental periodontal disease in rats monoinfected with gram-negative organisms. *J Periodontol Res*. 1978;13:326–32.
- Irving JT, Heeley JD, Amdur BH, Socransky SS. Pathological changes caused by the injection of lipids isolated from gram-positive organisms into the gingiva of rats. *J Periodontol Res*. 1979;14:160–6.
- Ismail MO, Greenman J, Morgan K, Glover MG, Rees AS, Scully C. Periodontitis in sheep: a model for human periodontal disease. *J Periodontol*. 1989;60:279–84.
- Iwatsuka H, Shino A, Suzuoki Z. General survey of diabetic features of yellow KK mice. *Endocrinol Jpn*. 1970;17:23–35.
- Jain A, Batista Jr EL, Serhan C, Stahl GL, Van Dyke TE. Role for periodontitis in the progression of lipid deposition in an animal model. *Infect Immun*. 2003;71:6012–8.
- Janssen U, Phillips AO, Floege J. Rodent models of nephropathy associated with type II diabetes. *J Nephrol*. 1999;12:159–72.
- Jensen SS, Broggin N, Weibrich G, Hjørtting-Hansen E, Schenk R, Buser D. Bone regeneration in standardized bone defects with autografts or bone substitutes in combination with platelet concentrate: a histologic and histomorphometric study in the mandibles of minipigs. *Int J Oral Maxillofac Implants*. 2005;20:703–12.
- Jensen S, Broggin N, Hjørtting-Hansen E, Schenk R, Buser D. Bone healing and graft resorption of autograft, anorganic bovine bone and b-tricalcium phosphate. A histologic and histomorphometric study in the mandibles of mini-pigs. *Clin Oral Implants Res*. 2006;17:237–43.
- Jogahara T, Oda S, Kawai T, Hanamura H, Koyasu K. Numerical variation of teeth in the wild house musk shrew *Suncus murinus* captured from Nagasaki, Japan. *Arch Oral Biol*. 2008;53:617–21.
- Jones WA, O'Leary TJ. The effectiveness of *in vivo* root planing in removing bacterial endotoxin from the roots of periodontally involved teeth. *J Periodontol*. 1978;49:337–42.
- Jordan HV, Keyes PH. Aerobic, gram-positive, filamentous bacteria as etiologic agents of experimental periodontal disease in hamsters. *Arch Oral Biol*. 1964;9:401–14.
- Jordan HV, Keyes PH, Bellack S. Periodontal lesions in hamsters and gnotobiotic rats infected with actinomycetes of human origin. *J Periodontol Res*. 1972;7:21–8.

- Jovanovic SA, Hunt DR, Bernard GW, Spiekermann H, Wozney JM, Wikesjö UM. Bone reconstruction following implantation of rhBMP-2 and guided bone regeneration in canine alveolar ridge defects. *Clin Oral Implants Res.* 2007;18:224–30.
- Jung UW, Chang YY, Um YJ, Kim CS, Cho KS, Choi SH. Interproximal periodontal defect model in dogs: a pilot study. *Oral Dis.* 2011;17:26–32.
- Karimbux NY, Ramamurthy NS, Golub LM, Nishimura I. The expression of collagen I and XII mRNAs in *Porphyromonas gingivalis*-induced periodontitis in rats: the effect of doxycycline and chemically modified tetracycline. *J Periodontol.* 1998;69:34–40.
- Kasiske BL, O'Donnell MP, Keane WF. The Zucker rat model of obesity, insulin resistance, hyperlipidemia, and renal injury. *Hypertension.* 1992;19 Suppl 1: 1110–5.
- Katz J, Ward DC, Michalek SM. Effect of host responses on the pathogenicity of strains of *Porphyromonas gingivalis*. *Oral Microbiol Immunol.* 1996;11:309–18.
- Katz J, Black KP, Michalek SM. Host responses to recombinant hemagglutinin B of *Porphyromonas gingivalis* in an experimental rat model. *Infect Immun.* 1999;67:4352–9.
- Kawai T, Eisen-Lev R, Seki M, Eastcott JW, Wilson ME, Taubman MA. Requirement of B7 costimulation for Th1-mediated inflammatory bone resorption in experimental periodontal disease. *J Immunol.* 2000;164: 2102–9.
- Kebschull M, Demmer RT, Papapanou PN. “Gum bug, leave my heart alone!” – epidemiologic and mechanistic evidence linking periodontal infections and atherosclerosis. *J Dent Res.* 2010;89:879–902.
- Kesavalu L, Chandrasekar B, Ebersole JL. In vivo induction of proinflammatory cytokines in mouse tissue by *Porphyromonas gingivalis* and *Actinobacillus actinomycetemcomitans*. *Oral Microbiol Immunol.* 2002;17: 177–80.
- Kesavalu L, Sathishkumar S, Bakthavatchalu V, Matthews C, Dawson D, Steffen M, Ebersole JL. Rat model of polymicrobial infection, immunity, and alveolar bone resorption in periodontal disease. *Infect Immun.* 2007;75:1704–12.
- Kiel RA, Kornman KS, Robertson PB. Clinical and microbiological effects of localized ligature-induced periodontitis on non-ligated sites in the cynomolgus monkey. *J Periodontol Res.* 1983;18:200–11.
- Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One.* 2009;4(11):e7824. doi:10.1371/journal.pone.0007824.
- Kim CK, Kim HY, Chai JK, et al. Effect of a calcium sulfate implant with calcium sulfate barrier on periodontal healing in 3-wall intrabony defects in dogs. *J Periodontol.* 1998;69:982–8.
- Kim HY, Kim CS, Jhon GJ, Moon IS, Choi SH, Cho KS, Chai JK, Kim CK. The effect of safflower seed extract on periodontal healing of 1-wall intrabony defects in beagle dogs. *J Periodontol.* 2002;73:1457–66.
- Kim CS, Choi SH, Chai JK, et al. Periodontal repair in surgically created intrabony defects in dogs: influence of the number of bone walls on healing response. *J Periodontol.* 2004;75:229–35.
- Kim CS, Choi SH, Cho KS, Chai JK, Wikesjö UM, Kim CK. Periodontal healing in one-wall intra-bony defects in dogs following implantation of autogenous bone or a coral-derived biomaterial. *J Clin Periodontol.* 2005; 32:583–9.
- Kim IY, Jung UW, Kim CS, Lee YK, Cho KS, Chai JK, Kim CK, Choi SH. Effects of a tetracycline blended polylactic and polyglycolic acid membrane on the healing of one-wall intrabony defects in beagle dogs. *Biomed Mater.* 2007;2:S106–10.
- Kim TG, Wikesjö UM, Cho KS, Chai JK, Pippig SD, Siedler M, Kim CK. Periodontal wound healing/regeneration following implantation of recombinant human growth/differentiation factor-5 (rhGDF-5) in an absorbable collagen sponge carrier into one-wall intrabony defects in dogs: a dose-range study. *J Clin Periodontol.* 2009;36:589–97.
- Kim CS, Um YJ, Chai JK, Cho KS, Moon IS, Choi SH, Jung UW, Lee DW, Kim CK. A canine model for histometric evaluation of periodontal regeneration. *Periodontol 2000.* 2011;56:209–26.
- King GN, King N, Cruchley AT, Wozney JM, Hughes FJ. Recombinant human bone morphogenetic protein-2 promotes wound healing in rat periodontal fenestration defects. *J Dent Res.* 1997;76:1460–70.
- King GN, King N, Hughes FJ. The effect of root surface demineralization on bone morphogenetic protein-2-induced healing of rat periodontal fenestration defects. *J Periodontol.* 1998;69:561–70.
- Kirkham J, Robinson C, Spence JA. Effect of periodontal disease (“broken mouth”) on the distribution of matrix macromolecules in the sheep periodontium. *Arch Oral Biol.* 1991;36:257–63.
- Kirkwood KL, Li F, Rogers JE, Otremba J, Coatney DD, Kreider JM, D'Silva NJ, Chakravarty S, Dugar S, Higgins LS, Protter AA, Medicherla S. A p38alpha selective mitogen-activated protein kinase inhibitor prevents periodontal bone loss. *J Pharmacol Exp Ther.* 2007;320:56–63.
- Klausen B. Microbiological and immunological aspects of experimental periodontal disease in rats: a review article. *J Periodontol.* 1991;62:59–73.
- Klausen B, Evans RT, Sfantescu C. Two complementary methods of assessing periodontal bone level in rats. *Scand J Dent Res.* 1989a;97:494–9.
- Klausen B, Hougen HP, Fiehn NE. Increased periodontal bone loss in temporarily B lymphocyte-deficient rats. *J Periodontol Res.* 1989b;24:384–90.
- Klepp M, Hinrichs JE, Eastlund T, Schaffer EM. Histologic evaluation of demineralized freeze-dried bone allografts in barrier membrane covered periodontal fenestration wounds and ectopic sites in dogs. *J Clin Periodontol.* 2004;31:534–44.
- Koo KT, Polimeni G, Albandar JM, Wikesjö UM. Periodontal repair in dogs: analysis of histometric assessments in the supraalveolar periodontal defect model. *J Periodontol.* 2004a;75:1688–93.

- Koo KT, Polimeni G, Albandar JM, Wikesjö UM. Periodontal repair in dogs: examiner reproducibility in the supraalveolar periodontal defect model. *J Clin Periodontol.* 2004b;31:439–42.
- Koo KT, Polimeni G, Qahash M, Kim CK, Wikesjö UM. Periodontal repair in dogs: guided tissue regeneration enhances bone formation in sites implanted with a coral-derived calcium carbonate biomaterial. *J Clin Periodontol.* 2005;32(1):104–10.
- Kornman KS, Holt SC, Roberston PB. The microbiology of ligature-induced periodontitis in the cynomolgus monkey. *J Periodontol Res.* 1981a;16:363–71.
- Kornman KS, Siegrist B, Soskolne WA, Nuki K. The predominant cultivable subgingival flora of beagle dogs following ligature placement and metronidazole therapy. *J Periodontol Res.* 1981b;16:251–8.
- Kornman KS, Manti F, Goldschneider I. Peripheral blood lymphocyte populations in ligature-induced periodontitis. *J Periodontol Res.* 1982;17:469–71.
- Kostopoulos L, Karring T. Susceptibility of GTR-regenerated periodontal attachment to ligature-induced periodontitis. *J Clin Periodontol.* 2004;31:336–40.
- Kozlovsky A, Tal H, Laufer BZ, Leshem R, Rohrer MD, Weinreb M, Artzi Z. Impact of implant overloading on the peri-implant bone in inflamed and non-inflamed peri-implant mucosa. *Clin Oral Implants Res.* 2007;18:601–10.
- Kuhr A, Popa-Wagner A, Schmoll H, Schwahn C, Kocher T. Observations on experimental marginal periodontitis in rats. *J Periodontol Res.* 2004;39:101–6.
- Kurtis B, Balos K, Oygür T. Effect of a collagen membrane enriched with fibronectin on guided tissue regeneration in dogs. *Periodontol Clin Investig.* 2002;24:11–9.
- Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *J Pharmacol Pharmacother.* 2010;1(2):94–9.
- Lalla E, Lamster IB, Feit M, Huang L, Schmidt AM. A murine model of accelerated periodontal disease in diabetes. *J Periodontol Res.* 1998;33:387–99.
- Lalla E, Lamster IB, Feit M, Huang L, Spessot A, Qu W, Kislinger T, Lu Y, Stern DM, Schmidt AM. Blockade of RAGE suppresses periodontitis-associated bone loss in diabetic mice. *J Clin Invest.* 2000;105:1117–24.
- Lalla E, Lamster IB, Hofmann MA, Bucciarelli L, Jerud AP, Tucker S, Lu Y, Papapanou PN, Schmidt AM. Oral infection with a periodontal pathogen accelerates early atherosclerosis in apolipoprotein E-null mice. *Arterioscler Thromb Vasc Biol.* 2003;23:1405–11.
- Lang H, Schüler N, Nolden R. Attachment formation following replantation of cultured cells into periodontal defects – a study in minipigs. *J Dent Res.* 1998;77:393–405.
- Le Guehennec L, Goyenvalle E, Aguado E, Houchmand-Cuny M, Enkel B, Pilet P, Daculsi G, Layrolle P. Small-animal models for testing macroporous ceramic bone substitutes. *J Biomed Mater Res B Appl Biomater.* 2005;72:69–78.
- Lee JY, Yi NN, Kim US, Choi JS, Kim SJ, Choi JI. *Porphyromonas gingivalis* heat shock protein vaccine reduces the alveolar bone loss induced by multiple periodontopathogenic bacteria. *J Periodontol Res.* 2006;41:10–4.
- Lee J, Tran Q, Seeba G, Wikesjö UM, Susin C. The critical-size supraalveolar peri-implant defect model: reproducibility in histometric data acquisition of alveolar bone formation and osseointegration. *J Clin Periodontol.* 2009;36:1067–74.
- Lekic P, Klausen B, Friis-Hasche E, Beloica D, Knezevic M, Hougen HP. Influence of age and immunization on development of gingivitis in rats. *Acta Odontol Scand.* 1989;47:233–8.
- Lekovic V, Klokkevold PR, Kenney EB, Dimitrijelic B, Nedic M, Weinlaender M. Histologic evaluation of guided tissue regeneration using 4 barrier membranes: a comparative furcation study in dogs. *J Periodontol.* 1998;69:54–61.
- Leone CW, Bokhadhoor H, Kuo D, Desta T, Yang J, Siqueira MF, Amar S, Graves DT. Immunization enhances inflammation and tissue destruction in response to *Porphyromonas gingivalis*. *Infect Immun.* 2006;74:2286–92.
- Li CH, Amar S. Morphometric, histomorphometric, and microcomputed tomographic analysis of periodontal inflammatory lesions in a murine model. *J Periodontol.* 2007;78(6):1120–8.
- Li L, Khansari A, Shapira L, Graves DT, Amar S. Contribution of interleukin-11 and prostaglandin(s) in lipopolysaccharide-induced bone resorption in vivo. *Infect Immun.* 2002a;70:3915–22.
- Li L, Messas E, Batista Jr EL, Levine RA, Amar S. *Porphyromonas gingivalis* infection accelerates the progression of atherosclerosis in a heterozygous apolipoprotein E-deficient murine model. *Circulation.* 2002b;105(7):861–7.
- Liang S, Hosur KB, Domon H, Hajishengallis G. Periodontal inflammation and bone loss in aged mice. *J Periodontol Res.* 2010;45:574–8.
- Liao F, Li Z, Wang Y, Shi B, Gong Z, Cheng X. *Porphyromonas gingivalis* may play an important role in the pathogenesis of periodontitis-associated rheumatoid arthritis. *Med Hypotheses.* 2009;72:732–5.
- Liberman DN, Pilau RM, Gaio EJ, Orlandini LF, Rösing CK. Low concentration alcohol intake may inhibit spontaneous alveolar bone loss in Wistar rats. *Arch Oral Biol.* 2011;56:109–13.
- Lin D, Smith MA, Champagne C, Elter J, Beck J, Offenbacher S. *Porphyromonas gingivalis* infection during pregnancy increases maternal tumor necrosis factor alpha, suppresses maternal interleukin-10, and enhances fetal growth restriction and resorption in mice. *Infect Immun.* 2003;71:5156–62.
- Lindhe J, Ericsson I. Effect of ligature placement and dental plaque on periodontal tissue breakdown in the dog. *J Periodontol.* 1978;49:343–50.
- Lindhe J, Rylander H. Experimental gingivitis in young dogs. *Scand J Dent Res.* 1975;83:314–26.

- Lindhe J, Hamp SE, Löe H. Plaque induced periodontal disease in beagle dogs. A 4-year clinical, roentgenographical and histometrical study. *J Periodontol Res*. 1975;10:243–55.
- Lindhe J, Berglundh T, Ericsson I, Liljenberg B, Marinello CP. Experimental breakdown of periimplant and periodontal tissues. A study in the beagle dog. *Clin Oral Implants Res*. 1992;3:9–16.
- Listgarten MA, Johnson D, Nowotny A, Tanner AC, Socransky SS. Histopathology of periodontal disease in gnotobiotic rats monoinfected with *Eikenella corrodens*. *J Periodontol Res*. 1978;13:134–48.
- Listgarten MA, Johnson D, Nowotny A, Tanner AC, Socransky SS. Histopathology of periodontal disease in gnotobiotic rats monoinfected with *Eikenella corrodens*. *J Periodontol Res*. 1978;13(2):134–48.
- Liu R, Bal HS, Desta T, Behl Y, Graves DT. Tumor necrosis factor- α mediates diabetes-enhanced apoptosis of matrix-producing cells and impairs diabetic healing. *Am J Pathol*. 2006;168:757–64.
- Liu Y, Zheng Y, Ding G, et al. Periodontal ligament stem cell-mediated treatment for periodontitis in miniature swine. *Stem Cells*. 2008;26:1065–73.
- Liu PF, Haake SK, Gallo RL, Huang CM. A novel vaccine targeting *Fusobacterium nucleatum* against abscesses and halitosis. *Vaccine*. 2009;27:1589–95.
- Liu PF, Shi W, Zhu W, Smith JW, Hsieh SL, Gallo RL, Huang CM. Vaccination targeting surface FomA of *Fusobacterium nucleatum* against bacterial co-aggregation: implication for treatment of periodontal infection and halitosis. *Vaccine*. 2010a;28(19):3496–505.
- Liu Y-F, Wu L-A, Wang J, Wen L-Y, Wang X-J. Micro-computerized tomography analysis of alveolar bone loss in ligature- and nicotine-induced experimental periodontitis in rats. *J Periodontol Res*. 2010b;45:714–9.
- Llavaneras A, Ramamurthy NS, Heikkilä P, Teronen O, Salo T, Rifkin BR, Ryan ME, Golub LM, Sorsa T. A combination of a chemically modified doxycycline and a bisphosphonate synergistically inhibits endotoxin-induced periodontal breakdown in rats. *J Periodontol*. 2001;72:1069–77.
- Lohinai Z, Benedek P, Feher E, Gyorfi A, Rosivall L, Fazekas A, Salzman AL, Szabo C. Protective effects of mercaptoethylguanidine, a selective inhibitor of inducible nitric oxide synthase, in ligature-induced periodontitis in the rat. *Br J Pharmacol*. 1998;123:353–60.
- Ma JL, Pan JL, Tan BS, Cui FZ. Determination of critical size defect of minipig mandible. *J Tissue Eng Regen Med*. 2009;3:615–22.
- Mahamed DA, Marleau A, Alnaeeli M, Singh B, Zhang X, Penninger JM, Teng YT. G(–) anaerobes-reactive CD4+ T-cells trigger RANKL-mediated enhanced alveolar bone loss in diabetic NOD mice. *Diabetes*. 2005;54:1477–86.
- Mallison 3rd SM, Szakal AK, Ranney RR, Tew JG. Antibody synthesis specific for nonoral antigens in inflamed gingiva. *Infect Immun*. 1988;56:823–30.
- Mallonee DH, Harvey CE, Venner M, Hammond BF. Bacteriology of periodontal disease in the cat. *Arch Oral Biol*. 1988;33:677–83.
- Mann PH, Harper DS, Regnier S. Reduction of calculus accumulation in domestic ferrets with two dentifrices containing pyrophosphate. *J Dent Res*. 1990;69:451–3.
- Manti F, Kornman K, Goldschneider I. Effects of an immunomodulating agent on peripheral blood lymphocytes and subgingival microflora in ligature-induced periodontitis. *Infect Immun*. 1984;45:172–9.
- Martí A, Marcos A, Martínez JA. Obesity and immune function relationships. *Obes Rev*. 2001;2:131–40.
- Matsson L, Attstrom R. Histologic characteristics of experimental gingivitis in the juvenile and adult beagle dog. *J Clin Periodontol*. 1979;6:334–50.
- McCourtie J, Poxton IR, Brown R, Whittaker CR, Spence JA, Aitchison GU. A longitudinal study of the cultivable subgingival anaerobic bacteria isolated from sheep during the development of broken mouth periodontitis. *J Med Microbiol*. 1990;31:275–83.
- Meng H, Xie H, Chen Z. Evaluation of ligature-induced periodontitis in minipig. *Zhonghua Kou Qiang Yi Xue Za Zhi*. 1996;31:333–6.
- Mignini LE, Khan KS. Methodological quality of systematic reviews of animal studies: a survey of reviews of basic research. *BMC Med Res Methodol*. 2006;6:10.
- Mikuls TR, Payne JB, Reinhardt RA, Thiele GM, Maziarz E, Cannella AC. Antibody responses to *Porphyromonas gingivalis* (*P. gingivalis*) in subjects with rheumatoid arthritis and periodontitis. *Int Immunopharmacol*. 2009;9:38–42.
- Min DH, Kim MJ, Yun JH, Kim CS, Lee YK, Choi SH, Kim KN, Kim CK. Effect of calcium phosphate glass scaffold with chitosan membrane on the healing of alveolar bone in 1-wall intrabony defect in beagle dogs. *Key Eng Mater*. 2005;284–286:851–4.
- Min C-K, Wikesjö UME, Park J-C, Chae G-J, Pippig SD, Bastone P, Kim C-S, Kim C-K. Wound healing/regeneration using recombinant human growth/ differentiation factor-5-inaninjectable poly-lactide-co-glycolide-acid composite carrier and a one-wall intra-bony defect model in dogs. *J Clin Periodontol*. 2011;38:261–8.
- Mitsuta T, Horiuchi H, Shinoda H. Effects of topical administration of clodronate on alveolar bone resorption in rats with experimental periodontitis. *J Periodontol*. 2002;73:479–86.
- Miyauchi M, Takata T, Ito H, Ogawa I, Kudo Y, Takekoshi T, Nikai H. Distribution of macrophage lineage cells in rat gingival tissue after topical application of lipopolysaccharide: an immunohistochemical study using monoclonal antibodies: OX6, ED1 and ED2. *J Periodontol Res*. 1998;33:345–51.
- Moher D, Schulz KF, Altman DG for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2011;357:1191–4.
- Monjo M, Rubert M, Wohlfahrt JC, Rønold HJ, Ellingsen JE, Lyngstadaas SP. In vivo performance of absorbable collagen sponges with rosuvastatin in critical-size cortical bone defects. *Acta Biomater*. 2010;6:1405–12.

- Moritz AJ, Cappelli D, Lantz MS, Holt SC, Ebersole JL. Immunization with *Porphyromonas gingivalis* cysteine protease: effects on experimental gingivitis and ligature-induced periodontitis in *Macaca fascicularis*. *J Periodontol*. 1998;69:686–97.
- Nagini S. Of humans and hamsters: the hamster buccal pouch carcinogenesis model as a paradigm for oral oncogenesis and chemoprevention. *Anticancer Agents Med Chem*. 2009;9:843–52.
- Nakajima K, Hamada N, Takahashi Y, Sasaguri K, Tsukinoki K, Umemoto T, Sato S. Restraint stress enhances alveolar bone loss in an experimental rat model. *J Periodontal Res*. 2006;41:527–34.
- Naguib G, Al-Mashat H, Desta T, Graves DT. Diabetes prolongs the inflammatory response to a bacterial stimulus through cytokine dysregulation. *J Invest Dermatol*. 2004;123(1):87–92.
- Nisengard R, Blann D, Zelonis L, McHenry K, Reynolds H, Zambon J. Effects of immunization with *B. macacae* on induced periodontitis—preliminary findings. *Immunol Invest*. 1989;18:225–37.
- Nishimura M. Breeding of mice strains for diabetes mellitus. *Exp Anim*. 1969;18:147–57.
- Nkenke E, Lehner B, Weinzierl K, et al. Bone contact, growth, and density around immediately loaded implants in the mandible of minipigs. *Clin Oral Implants Res*. 2003;14:312–21.
- Nkenke E, Lehner B, Fenner M, et al. Immediate versus delayed loading of dental implants in the maxillae of minipigs: follow-up of implant stability and implant failures. *Int J Oral Maxillofac Implants*. 2005;20:39–47.
- Nociti Jr FH, Nogueira-Filho GR, Primo MT, Machado MA, Tramontina VA, Barros SP, et al. The influence of nicotine on the bone loss rate in ligature-induced periodontitis. A histometric study in rats. *J Periodontol*. 2000;71:1460–4.
- Nogueira-Filho GR, Froes Neto EB, Casati MZ, Reis SR, Tunes RS, Tunes UR. Nicotine effects on alveolar bone changes induced by occlusal trauma: a histometric study in rats. *J Periodontol*. 2004;75:348–52.
- Nowotny A, Sanavi F. Induction of nonspecific tolerance to endotoxins reduces the alveolar bone resorption in ligature-treated rats. *Infect Immun*. 1983;39:873–8.
- Núñez J, Caffesse R, Vignoletti F, Guerra F, San Roman F, Sanz M. Clinical and histological evaluation of an acellular dermal matrix allograft in combination with the coronally advanced flap in the treatment of Miller class I recession defects: an experimental study in the mini-pig. *J Clin Periodontol*. 2009;36:523–31.
- Núñez J, Sanz-Blasco S, Vignoletti F, Muñoz F, Caffesse RG, Sanz M, Villalobos C, Núñez L. 17beta-estradiol promotes cementoblast proliferation and cementum formation in experimental periodontitis. *J Periodontol*. 2010;81:1064–74.
- O'Brien-Simpson NM, Paolini RA, Reynolds EC. RgpA-Kgp peptide-based immunogens provide protection against *Porphyromonas gingivalis* challenge in a murine lesion model. *Infect Immun*. 2000;68:4055–63.
- O'Brien-Simpson NM, Pathirana RD, Paolini RA, Chen YY, Veith PD, Tam V, Ally N, Pike RN, Reynolds EC. An immune response directed to proteinase and adhesion functional epitopes protects against *Porphyromonas gingivalis*-induced periodontal bone loss. *J Immunol*. 2005;175:3980–9.
- Oates TW, Graves DT, Cochran DL. Clinical, radiographic and biochemical assessment of IL-1/TNF-alpha antagonist inhibition of bone loss in experimental periodontitis. *J Clin Periodontol*. 2002;29:137–43.
- Ohnishi T, Bandow K, Kakimoto K, Machigashira M, Matsuyama T, Matsuguchi T. Oxidative stress causes alveolar bone loss in metabolic syndrome model mice with type 2 diabetes. *J Periodontal Res*. 2009;44:43–51.
- Okada Y, Hamada N, Kim Y, Takahashi Y, Sasaguri K, Ozono S, Sato S. Blockade of sympathetic b-receptors inhibits *Porphyromonas gingivalis*-induced alveolar bone loss in an experimental rat periodontitis model. *Arch Oral Biol*. 2010;55:502–8.
- Onodera H, Shibukawa Y, Sugito H, Ota M, Yamada S. Periodontal regeneration in intrabony defects after application of enamel matrix proteins with guided tissue regeneration: an experimental study in dogs. *Biomed Res*. 2005;26:69–77.
- Oortgiesen DA, Meijer GJ, Bronckers AL, Walboomers XF, Jansen JA. Fenestration defects in the rabbit jaw: an inadequate model for studying periodontal regeneration. *Tissue Eng Part C Methods*. 2010;16: 133–40.
- Oz HS, Ebersole JL. A novel murine model for chronic inflammatory alveolar bone loss. *J Periodontal Res*. 2010;45:94–9.
- Oz HS, Chen T, Ebersole JL. A model for chronic mucosal inflammation in IBD and periodontitis. *Dig Dis Sci*. 2010;55:2194–202.
- Oz HS, Puleo DA. Animal models for periodontal disease. *J Biomed Biotechnol*. 2011:754–857.
- Ozmeric N, Bal B, Oygur T, Balos K. The effect of a collagen membrane in regenerative therapy of two-wall intrabony defects in dogs. *Periodontal Clin Invest*. 2000;22:22–30.
- Papageorgiou A, Vouros I, Konstantinidis A. Treatment outcomes of ligature-induced recession in the dog model using guided tissue regeneration or coronally positioned flap procedures. *J Int Acad Periodontol*. 2009;11:177–87.
- Park WY, Cho CS, Chai JK, Kim CK, Choi SH. The effects of bioactive glass on the periodontal healing of 1-wall intrabony defects in dogs. *J Korean Acad Periodontol*. 1998;28:145–60.
- Park JS, Choi SH, Moon IS, Cho KS, Chai JK, Kim CK. Eight-week histological analysis on the effect of chitosan on surgically created one-wall intrabony defects in beagle dogs. *J Clin Periodontol*. 2003;30:443–53.
- Park CH, Abramson ZR, Taba Jr M, Jin Q, Chang J, Kreider JM, Goldstein SA, Giannobile WV. Three-dimensional micro-computed tomographic imaging of alveolar bone in experimental bone loss or repair. *J Periodontol*. 2007;78:273–81.

- Pathirana RD, O'Brien-Simpson NM, Brammar GC, Slakeski N, Reynolds EC. Kgp and RgpB, but not RgpA, are important for *Porphyromonas gingivalis* virulence in the murine periodontitis model. *Infect Immun*. 2007;75:1436–42.
- Pellegrini G, Seol YJ, Gruber R, Giannobile WV. Pre-clinical models for oral and periodontal reconstructive therapies. *J Dent Res*. 2009;88:1065–76.
- Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, Macleod M, Mignini LE, Jayaram P, Khan KS. Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ*. 2007;334(7586):197.
- Perlstein MI, Bissada NF. Influence of obesity and hypertension on the severity of periodontitis in rats. *Oral Surg Oral Med Oral Pathol*. 1977;43:707–19.
- Persson GR. Immune responses and vaccination against periodontal infections. *J Clin Periodontol*. 2005;32 Suppl 6:39–53.
- Persson GR, Engel D, Whitney C, Darveau R, Weinberg A, Brunsvold M, Page RC. Immunization against *Porphyromonas gingivalis* inhibits progression of experimental periodontitis in nonhuman primates. *Infect Immun*. 1994a;62:1026–31.
- Persson GR, Engel LD, Whitney CW, Weinberg A, Moncla BJ, Darveau RP, Houston L, Braham P, Page RC. Macaca fascicularis as a model in which to assess the safety and efficacy of a vaccine for periodontitis. *Oral Microbiol Immunol*. 1994b;9:104–11.
- Persson LG, Berglundh T, Sennerby L, Lindhe J. Reosseointegration after treatment of peri-implantitis at different implant surfaces. An experimental study in the dog. *Clin Oral Implants Res*. 2001a;12:595–603.
- Persson LG, Ericsson I, Berglundh T, Lindhe J. Osseointegration following treatment of peri-implantitis and replacement of implant components. An experimental study in the dog. *J Clin Periodontol*. 2001b;28:258–63.
- Peruzzo DC, Benatti BB, Antunes IB, Andersen ML, Sallum EA, Casati MZ, et al. Chronic stress may modulate periodontal disease: a study in rats. *J Periodontol*. 2008;79:697–704.
- Pieri F, Lucarelli E, Corinaldesi G, Fini M, Aldini NN, Giardino R, Donati D, Marchetti C. Effect of mesenchymal stem cells and platelet-rich plasma on the healing of standardized bone defects in the alveolar ridge: a comparative histomorphometric study in minipigs. *J Oral Maxillofac Surg*. 2009;67:265–72.
- Pimentel SP, Sallum AW, Saldanha JB, Casati MZ, Nociti Jr FH, Sallum EA. Enamel matrix derivative versus guided tissue regeneration in the presence of nicotine: a histomorphometric study in dogs. *J Clin Periodontol*. 2006;33:900–7.
- Polak D, Wilensky A, Shapira L, Halabi A, Goldstein D, Weiss EI, Houry-Haddad Y. Mouse model of experimental periodontitis induced by *Porphyromonas gingivalis*/*Fusobacterium nucleatum* infection: bone loss and host response. *J Clin Periodontol*. 2009;36:406–10.
- Polak D, Wilensky A, Shapira L, Weiss EI, Houry-Haddad Y. Vaccination of mice with *Porphyromonas gingivalis* or *Fusobacterium nucleatum* modulates the inflammatory response, but fails to prevent experimental periodontitis. *J Clin Periodontol*. 2010;37:812–7.
- Polimeni G, Albandar JM, Wikesjö UME. Prognostic factors for alveolar regeneration: osteogenic potential of resident bone. *J Clin Periodontol*. 2004a;31:840–4.
- Polimeni G, Koo K-T, Qahash M, Xiropaidis AV, Albandar JM, Wikesjö UME. Prognostic factors for alveolar regeneration: bone formation at teeth and titanium implants. *J Clin Periodontol*. 2004b;31:927–32.
- Polimeni G, Koo K-T, Qahash M, Xiropaidis AV, Albandar JM, Wikesjö UME. Prognostic factors for alveolar regeneration: effect of tissue occlusion on alveolar bone regeneration with guided tissue regeneration. *J Clin Periodontol*. 2004c;31:730–5.
- Polimeni G, Koo K-T, Qahash M, Xiropaidis AV, Albandar JM, Wikesjö UME. Prognostic factors for alveolar regeneration: effect of a space-providing biomaterial on guided tissue regeneration. *J Clin Periodontol*. 2004d;31:725–9.
- Polimeni G, Albandar JM, Wikesjö UME. Prognostic factors for alveolar regeneration: effect of space provision. *J Clin Periodontol*. 2005;32:951–4.
- Polimeni G, Susin C, Wikesjö UME. Regenerative potential and healing dynamics of the periodontium: a critical-size supra-alveolar periodontal defect study. *J Clin Periodontol*. 2009;36:258–64.
- Polimeni G, Xiropaidis AV, Wikesjö UM. Biology and principles of periodontal wound healing/regeneration. *Periodontol* 2000. 2006;41:30–47.
- Pontes Andersen CC, Flyvbjerg A, Buschard K, Holmstrup P. Relationship between periodontitis and diabetes: lessons from rodent studies. *J Periodontol*. 2007;78:1264–75.
- Pouliot M, Clish CB, Petasis NA, Van Dyke TE, Serhan CN. Lipoxin A(4) analogues inhibit leukocyte recruitment to *Porphyromonas gingivalis*: a role for cyclooxygenase-2 and lipoxins in periodontal disease. *Biochemistry*. 2000;39:4761–8.
- Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I. Reviewing Animal Trials Systematically (RATS) Group. Where is the evidence that animal research benefits humans? *BMJ*. 2004;328(7438):514–7.
- Powell CA, Bannister SR, Mackey SA, Maller SC, McDonnell HT, Deas DE. Periodontal wound healing with and without platelet-rich plasma: histologic observations and assessment of flap tensile strength. *J Periodontol*. 2009;80:985–92.
- Rajapakse PS, O'Brien-Simpson NM, Slakeski N, Hoffmann B, Reynolds EC. Immunization with the RgpA-Kgp proteinase-adhesin complexes of *Porphyromonas gingivalis* protects against periodontal bone loss in the rat periodontitis model. *Infect Immun*. 2002;70:2480–6.
- Romanowski AW, Squier CA, Lesch CA. Permeability of rodent junctional epithelium to exogenous protein. *J Periodontal Res*. 1988;23:81–6.
- Ramamurthy NS, Rifkin BR, Greenwald RA, Xu JW, Liu Y, Turner G, Golub LM, Vernillo AT. Inhibition of matrix metalloproteinase-mediated periodontal

- bone loss in rats: a comparison of 6 chemically modified tetracyclines. *J Periodontol.* 2002a;73:726–34.
- Ramamurthy NS, Xu JW, Bird J, Baxter A, Bhogal R, Wills R, Watson B, Owen D, Wolff M, Greenwald RA. Inhibition of alveolar bone loss by matrix metalloproteinase inhibitors in experimental periodontal disease. *J Periodontol Res.* 2002b;37:1–7.
- Ramamurthy NS, Greenwald RA, Celiker MY, Shi EY. Experimental arthritis in rats induces biomarkers of periodontitis which are ameliorated by gene therapy with tissue inhibitor of matrix metalloproteinases. *J Periodontol.* 2005;76:229–33.
- Reichart PA, Dürr UM, Triadan H, Vickendey G. Periodontal disease in the domestic cat. A histopathologic study. *J Periodontol Res.* 1984;19:67–75.
- Repeke CE, Ferreira Jr SB, Claudino M, Silveira EM, de Assis GF, Avila-Campos MJ, Silva JS, Garlet GP. Evidences of the cooperative role of the chemokines CCL3, CCL4 and CCL5 and its receptors CCR1+ and CCR5+ in RANKL+ cell migration throughout experimental periodontitis in mice. *Bone.* 2010;46:1122–30.
- Reymond F, Cimasoni G. Radiographic measurement of alveolar bone loss in the living rat. *J Periodontol.* 1997;68:141–4.
- Rizzo AA, Mitchell CT. Chronic allergic inflammation induced by repeated deposition of antigen in rabbit gingival pockets. *Periodontics.* 1966;4:5–10.
- Rogers JE, Li F, Coatney DD, Rossa C, Bronson P, Krieder JM, Giannobile WV, Kirkwood KL. *Actinobacillus actinomycetemcomitans* lipopolysaccharide-mediated experimental bone loss model for aggressive periodontitis. *J Periodontol.* 2007;78:550–8.
- Roriz VM, Souza SL, Taba Jr M, Palioto DB, Grisi MF. Treatment of class III furcation defects with expanded polytetrafluoroethylene membrane associated or not with anorganic bone matrix/synthetic cell-binding peptide: a histologic and histomorphometric study in dogs. *J Periodontol.* 2006;77:490–7.
- Ross BC, Czajkowski L, Vandenberg KL, Camuglia S, Woods J, Agius C, Paolini R, Reynolds E, Barr IG. Characterization of two outer membrane protein antigens of *Porphyromonas gingivalis* that are protective in a murine lesion model. *Oral Microbiol Immunol.* 2004;19:6–15.
- Ryder MI. Histological and ultrastructural characteristics of the periodontal syndrome in the rice rat. I. General light microscopic observations and ultrastructural observations of initial inflammatory changes. *J Periodontol Res.* 1980;15:502–15.
- Sallum EA, Pimentel SP, Saldanha JB, Nogueira-Filho GR, Casati MZ, Nociti FH, Sallum AW. Enamel matrix derivative and guided tissue regeneration in the treatment of dehiscence-type defects: a histomorphometric study in dogs. *J Periodontol.* 2004;75(10):1357–63.
- Sallay K, Sanavi F, Ring I, Pham P, Behling UH, Nowotny A. Alveolar bone destruction in the immunosuppressed rat. *J Periodontol Res.* 1982;17(3):263–74.
- Saito E, Saito A, Kawanami M. Favorable healing following space creation in rhBMP-2-induced periodontal regeneration of horizontal circumferential defects in dogs with experimental periodontitis. *J Periodontol.* 2003;74:1808–15.
- Sakata J, Abe H, Ohazama A, et al. Effects of combined treatment with porous bovine inorganic bone grafts and bilayer porcine collagen membrane on refractory one-wall intrabony defects. *Int J Periodontics Restorative Dent.* 2006;26:161–9.
- Sasaki H, Hou L, Belani A, Wang CY, Uchiyama T, Müller R, Stashenko P. IL-10, but not IL-4, suppresses infection-stimulated bone resorption in vivo. *J Immunol.* 2000;165:3626–30.
- Sasaki H, Suzuki N, AlShwaimi E, Xu Y, Battaglini R, Morse L, Stashenko P. 18b-Glycyrrhetic acid inhibits periodontitis via glucocorticoid-independent nuclear factor- κ B inactivation in interleukin-10-deficient mice. *J Periodontol Res.* 2010;45:757–63.
- Schenkein HA. Failure of *Bacteroides gingivalis* W83 to accumulate bound C3 following opsonization with serum. *J Periodontol Res.* 1989;24:20–7.
- Schliephake H, Alejt J. Mandibular onlay grafting using prefabricated bone grafts with primary implant placement: an experimental study in minipigs. *Int J Oral Maxillofac Implants.* 1998;13:384–93.
- Schmitt JM, Buck DC, Joh SP, Lynch SE, Hollinger JO. Comparison of porous bone mineral and biologically active glass in critical-sized defects. *J Periodontol.* 1997;68:1043–53.
- Schou S, Holmstrup P, Kornman KS. Non-human primates used in studies of periodontal disease pathogenesis: a review of the literature. *J Periodontol.* 1993;64:497–508.
- Schou S, Holmstrup P, Stoltze K, Hjørting-Hansen E, Fiehn NE, Skovgaard LT. Probing around implants and teeth with healthy or inflamed peri-implant mucosa/gingiva. *Clin Oral Implants Res.* 2002;13:113–26.
- Schulz KF, Altman DG, Moher D, The CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c332.
- Schwarz F, Herten M, Sager M, Bieling K, Sculean A, Becker J. Comparison of naturally occurring and ligature-induced periimplantitis bone defects in humans and dogs. *Clin Oral Implants Res.* 2007;18:161–70.
- Sculean A, Karring T, Theilade J, Lioubavina N. The regenerative potential of oxytalan fibers: an experimental study in the monkey. *J Clin Periodontol.* 1997;24:932–6.
- Sculean A, Donos N, Reich E, Karring T, Brex M. Regeneration of oxytalan fibres in different types of periodontal defects: a histological study in monkeys. *J Periodontol Res.* 1998;33:453–9.
- Sculean A, Donos N, Brex M, Karring T, Reich E. Healing of fenestration-type defects following treatment with guided tissue regeneration or enamel matrix proteins: an experimental study in monkeys. *Clin Oral Invest.* 2000a;4:50–6.

- Sculean A, Donos N, Brex M, Reich E, Karring T. Treatment of intrabony defects with guided tissue regeneration and enamel matrix-proteins: an experimental study in monkeys. *J Clin Periodontol*. 2000b;27:466–72.
- Sculean A, Berakdar M, Pahl S, Windisch P, Brex M, Reich E, Donos N. Patterns of cytokeratin expression in monkey and human periodontium following regenerative and conventional periodontal surgery. *J Periodontol Res*. 2001;36:260–8.
- Sculean A, Nikolidakis D, Schwarz F. Regeneration of periodontal tissues: combinations of barrier membranes and grafting materials – biological foundation and preclinical evidence: a systematic review. *J Clin Periodontol*. 2008;35(8 Suppl):106–16.
- Selvig KA, Wikesjö UM, Bogle GC, Finkelman RD. Impaired early bone formation in periodontal fenestration defects in dogs following application of insulin-like growth factor (II). Basic fibroblast growth factor and transforming growth factor beta 1. *J Clin Periodontol*. 1994;21:380–5.
- Semenoff Segundo A, Semenoff TA, Borges AH, Pedro FL, Sakai VT. Methodological model of chronic stress associated with ligature-induced periodontitis in rats: a radiographic study. *Braz Oral Res*. 2010;24:455–9.
- Serhan CN, Jain A, Marleau S, et al. Reduced inflammation and tissue damage in transgenic rabbits overexpressing 15-lipoxygenase and endogenous anti-inflammatory lipid mediators. *J Immunol*. 2003;171:6856–65.
- Seto H, Toba Y, Takada Y, Kawakami H, Ohba H, Hama H, Horibe M, Nagata T. Milk basic protein increases alveolar bone formation in rat experimental periodontitis. *J Periodontol Res*. 2007;42:85–9.
- Sharma A, Honma K, Evans RT, Hruby DE, Genco RJ. Oral immunization with recombinant *Streptococcus gordonii* expressing *porphyromonas gingivalis* FimA domains. *Infect Immun*. 2001;69:2928–34.
- Sharma A, Inagaki S, Honma K, Sfintescu C, Baker PJ, Evans RT. Tannerella forsythia-induced alveolar bone loss in mice involves leucine-rich-repeat BspA protein. *J Dent Res*. 2005;84:462–7.
- Shirakata Y, Oda S, Kinoshita A, Kikuchi S, Tsuchioka H, Ishikawa I. Histocompatible healing of periodontal defects after application of an injectable calcium phosphate bone cement. A preliminary study in dogs. *J Periodontol*. 2002;73:1043–53.
- Shirakata Y, Yoshimoto T, Goto H, Yonamine Y, Kadomatsu H, Miyamoto M, Nakamura T, Hayashi C, Izumi Y. Favorable periodontal healing of 1-wall intrabony defects after application of calcium phosphate cement wall alone or in combination with enamel matrix derivative: a pilot study with canine mandibles. *J Periodontol*. 2007;78:889–98.
- Shoji K, Horiuchi H, Shinoda H. Inhibitory effects of a bisphosphonate (risedronate) on experimental periodontitis in rats. *J Periodontol Res*. 1995;30:277–84.
- Shoji K, Ohtsuka-Isaya M, Shimauchi H, Shinoda H. Effects of lactation on alveolar bone loss in experimental periodontitis. *J Periodontol*. 2007;78:152–6.
- Siegrist B, Korman KS. The effect of supragingival plaque control on the composition of the subgingival microbial flora in ligature-induced periodontitis in the monkey. *J Dent Res*. 1982;61:936–41.
- Singh G, O'Neal RB, Brennan WA, Strong SL, Horner JA, Van Dyke TE. Surgical treatment of induced periodontitis in the micro pig: clinical and histological analysis. *J Periodontol*. 1993;64:984–9.
- Sluyter F, Breivik T, Cools A. Manipulations in maternal environment reverse periodontitis in genetically predisposed rats. *Clin Diagn Lab Immunol*. 2002;9:931–2.
- Soames JV, Davies RM. Lymphocyte-macrophage and lymphocyte-lymphocyte associations in early gingivitis in beagle dogs. *J Periodontol Res*. 1980;15:341–4.
- Socransky SS, Hubersak C, Propas D. Induction of periodontal destruction in gnotobiotic rats by a human oral strain of *Actinomyces naeslundii*. *Arch Oral Biol*. 1970;15:993–5.
- Soletti AC, Gaio EJ, Rosing CK. Effect of neonatal clomipramine in the pathogenesis of ligature-induced periodontitis in Lewis rats. *Acta Odontol Scand*. 2009;67:94–8.
- Song WS, Kim CS, Choi SH, Jhon GJ, Kim HY, Cho KS, Kim CK, Chai JK. The effects of a bioabsorbable barrier membrane containing safflower seed extracts on periodontal healing of 1-wall intrabony defects in beagle dogs. *J Periodontol*. 2005;76:22–33.
- Stavropoulos A, Wikesjö UME. Influence of defect dimensions on periodontal wound healing/regeneration in intrabony defects following implantation of a bovine bone biomaterial and provisions for guided tissue regeneration: an experimental study in the dog. *J Clin Periodontol*. 2010;37:534–43.
- Struillou X, Boutigny H, Soueidan A, Layrolle P. Experimental animal models in periodontology: a review. *Open Dent J*. 2010;4:37–47.
- Suaid FF, Ribeiro FV, Rodrigues TL, Silvério KG, Carvalho MD, Nociti FH Jr, Casati MZ, Sallum EA. Autologous periodontal ligament cells in the treatment of class II furcation defects: a study in dogs. *J Clin Periodontol*. 2011;38:491–8.
- Suzumura Y, Kameyama Y, Mizutani M, Kato M, Kondo K, Mabuchi R. Long junctional epithelium produced by application of bacterial protease in rats. *J Periodontol Res*. 1989;24:217–21.
- Takahashi D, Odajima T, Morita M, Kawanami M, Kato H. Formation and resolution of ankylosis under application of recombinant human bone morphogenetic protein-2 (rhBMP-2) to class III furcation defects in cats. *J Periodontol Res*. 2005;40:299–305.
- Takano M, Nishihara R, Sugano N, Matsumoto K, Yamada Y, Takane M, Fujisaki Y, Ito K. The effect of systemic anti-tumor necrosis factor-alpha treatment on *Porphyromonas gingivalis* infection in type 2 diabetic mice. *Arch Oral Biol*. 2010;55:379–84.
- Takata T, Matsuura M, Murashima M, Miyauchi M, Nikai H. Periodontitis in the house musk shrew (*Suncus murinus*): a potential animal model for human periodontal disease. *J Periodontol*. 1999;70:195–200.
- Tal H, Pitaru S, Moses O, Kozlovsky A. Collagen gel and membrane in guided tissue regeneration in periodontal fenestration defects in dogs. *J Clin Periodontol*. 1996;23:1–6.

- Tal H, Artzi Z, Moses O, Nemcovsky C, Kozlovsky A. Guided periodontal regeneration using bilayered collagen membranes and bovine bone mineral in fenestration defects in the canine. *Int J Periodontics Restorative Dent*. 2005;25:509–18.
- Teng YT, Nguyen H, Gao X, Kong YY, Gorczynski RM, Singh B, Ellen RP, Penninger JM. Functional human T-cell immunity and osteoprotegerin ligand control alveolar bone destruction in periodontal infection. *J Clin Invest*. 2000;106(6):R59–67.
- Teng YT, Nguyen H, Hassanloo A, Ellen RP, Hozumi N, Gorczynski RM. Periodontal immune responses of human lymphocytes in *Actinobacillus actinomycetemcomitans*-inoculated NOD/SCID mice engrafted with peripheral blood leukocytes of periodontitis patients. *J Periodontol Res*. 1999;34(1):54–61.
- Teng YT. Mixed periodontal Th1-Th2 cytokine profile in *Actinobacillus actinomycetemcomitans*-specific osteoprotegerin ligand (or RANK-L)- mediated alveolar bone destruction in vivo. *Infect Immun*. 2002;70:5269–73.
- Teng YT. The role of acquired immunity and periodontal disease progression. *Crit Rev Oral Biol Med*. 2003;14:237–52.
- Teng YT, Hu W. Expression cloning of a periodontitis-associated apoptotic effector, cagE homologue, in *Actinobacillus actinomycetemcomitans*. *Biochem Biophys Res Commun*. 2003;303:1086–94.
- Teng YT, Zhang X. Apoptotic activity and sub-cellular localization of a T4SS-associated CagE-homologue in *Actinobacillus actinomycetemcomitans*. *Microb Pathog*. 2005;38:125–32.
- Teng YT, Mahamed D, Singh B. Gamma interferon positively modulates *Actinobacillus actinomycetemcomitans*-specific RANKL+ CD4+ Th-cell-mediated alveolar bone destruction in vivo. *Infect Immun*. 2005;73:3453–61.
- Thoma DS, Halg GA, Dard MM, Seibl R, Hammerle CH, Jung RE. Evaluation of a new biodegradable membrane to prevent gingival ingrowth into mandibular bone defects in minipigs. *Clin Oral Implants Res*. 2009;20:7–16.
- Toker H, Ozan F, Ozer H, Ozdemir H, Eren K, Yeler H. A morphometric and histopathologic evaluation of the effects of propolis on alveolar bone loss in experimental periodontitis in rats. *J Periodontol*. 2008;79(6):1089–94.
- Tomita S, Yamamoto S, Shibukawa Y, Kaneko T, Miyakoshi S, Shimono M, Yamada S. Application of 4-META/MMA-TBB resin for fixation of membrane to tooth in guided tissue regeneration in dog. *Dent Mater J*. 2010;29:690–6.
- Tomofuji T, Ekuni D, Yamanaka R, Kusano H, Azuma T, Sanbe T, et al. Chronic administration of lipopolysaccharide and proteases induces periodontal inflammation and hepatic steatosis in rats. *J Periodontol*. 2007;78:1999–2006.
- Tomofuji T, Sanbe T, Ekuni D, Azuma T, Irie K, Maruyama T, Tamaki N, Yamamoto T. Oxidative damage of rat liver induced by ligature-induced periodontitis and chronic ethanol consumption. *Arch Oral Biol*. 2008;53:1113–8.
- Tomofuji T, Ekuni D, Irie K, Azuma T, Endo Y, Tamaki N, Sanbe T, Murakami J, Yamamoto T, Morita M. Preventive effects of a cocoa-enriched diet on gingival oxidative stress in experimental periodontitis. *J Periodontol*. 2009a;80:1799–808.
- Tomofuji T, Ekuni D, Sanbe T, Azuma T, Tamaki N, Irie K, Maruyama T, Yamamoto T, Watanabe T, Miyauchi M, Takata T. Effects of improvement in periodontal inflammation by toothbrushing on serum lipopolysaccharide concentration and liver injury in rats. *Acta Odontol Scand*. 2009b;67:1–6.
- Tomofuji T, Ekuni D, Sanbe T, et al. Effects of improvement in periodontal inflammation by toothbrushing on serum lipopolysaccharide concentration and liver injury in rats. *Acta Odontol Scand*. 2009c;67:200–5.
- Tomofuji T, Yamamoto T, Tamaki N, Ekuni D, Azuma T, Sanbe T, Irie K, Kasuyama K, Umakoshi M, Murakami J, Koikeguchi S, Morita M. Effects of obesity on gingival oxidative stress in a rat model. *J Periodontol*. 2009d;80:1324–9.
- Tyrrell KL, Citron DM, Jenkins JR, Goldstein EJ. Periodontal bacteria in rabbit mandibular and maxillary abscesses. *J Clin Microbiol*. 2002;40:1044–7.
- Vardar-Sengul S, Buduneli E, Turkoglu O, Buduneli N, Atilla G, Wahlgren J, Sorsa T, Baylas H. The effects of selective COX-2 inhibitor/celecoxib and omega-3 fatty acid on matrix metalloproteinases, TIMP-1, and laminin-5gamma2-chain immunolocalization in experimental periodontitis. *J Periodontol*. 2008;79:1934–41.
- Vastardis S, Yukna RA, Mayer ET, Atkinson BL. Periodontal regeneration with peptide-enhanced anorganic bone matrix in particulate and putty form in dogs. *J Periodontol*. 2005;76:1690–6.
- Verma RK, Bhattacharyya I, Sevilla A, Lieberman I, Pola S, Nair M, Wallet SM, Aukhil I, Kesavalu L. Virulence of major periodontal pathogens and lack of humoral immune protection in a rat model of periodontal disease. *Oral Dis*. 2010;16:686–95.
- Villaca JH, Novaes Jr AB, Souza SL, Taba Jr M, Molina GO, Carvalho TL. Bioactive glass efficacy in the periodontal healing of intrabony defects in monkeys. *Braz Dent J*. 2005;16:67–74.
- Wang HL, Pappert TD, Castelli WA, Chiego Jr DJ, Shyr Y, Smith BA. The effect of platelet-derived growth factor on the cellular response of the periodontium: an autoradiographic study on dogs. *J Periodontol*. 1994;65:429–36.
- Wang M, Shakhathreh MA, James D, Liang S, Nishiyama S, Yoshimura F, Demuth DR, Hajishengallis G. Fimbrial proteins of *porphyromonas gingivalis* mediate in vivo virulence and exploit TLR2 and complement receptor 3 to persist in macrophages. *J Immunol*. 2007a;179:2349–58.
- Wang S, Liu Y, Fang D, Shi S. The miniature pig: a useful large animal model for dental and orofacial research. *Oral Dis*. 2007b;13:530–7.
- Watanabe K, Petro BJ, Shlimon AE, Unterman TG. Effect of periodontitis on insulin resistance and the onset of type 2 diabetes mellitus in Zucker diabetic fatty rats. *J Periodontol*. 2008;79:1208–16.

- Watanabe K, Iizuka T, Adeleke A, Pham L, Shlimon AE, Yasin M, Horvath P, Unterman TG. Involvement of toll-like receptor 4 in alveolar bone loss and glucose homeostasis in experimental periodontitis. *J Periodontol Res*. 2011;46(1):21–30.
- Weinberg MA, Bral M. Laboratory animals models in periodontology. *J Clin Periodontol*. 1999;26:335–40.
- Weng D, Hürzeler MB, Quinones CR, Pechstadt B, Mota L, Caffesse RG. Healing patterns in recession defects treated with ePTFE membranes and with free connective tissue grafts: a histologic and histometric study in the beagle dog. *J Clin Periodontol*. 1998;25:238–45.
- Wikesjö UM, Nilvéus R. Periodontal repair in dogs: effect of wound stabilization on healing. *J Periodontol*. 1990;61:719–24.
- Wikesjö UM, Selvig KA. Periodontal wound healing and regeneration. *Periodontol* 2000. 1999;19:21–39.
- Wikesjö UM, Hagen K, Nielsen DD. Periodontal repair in dogs: effect of saliva contamination of the root surface. *J Periodontol*. 1990;61:559–63.
- Wikesjö UM, Selvig KA, Zimmerman G, Nilvéus R. Periodontal repair in dogs: healing in experimentally created chronic periodontal defects. *J Periodontol*. 1991;62:258–63.
- Wikesjö UME, Kean CJC, Zimmerman GJ. Periodontal repair in dogs: supraalveolar defect models for evaluation of safety and efficacy of periodontal reconstructive therapy. *J Periodontol*. 1994;65:1151–7.
- Wikesjö UM, Xirapaidis AV, Thomson RC, Cook AD, Selvig KA, Hardwick WR. Periodontal repair in dogs: rhBMP-2 significantly enhances bone formation under provisions for guided tissue regeneration. *J Clin Periodontol*. 2003a;30:705–14.
- Wikesjö UME, Lim WH, Thomson RC, Hardwick WR. Periodontal repair in dogs: gingival tissue occlusion, a critical requirement for GTR? *J Clin Periodontol*. 2003b;30:655–64.
- Wikesjö UME, Susin C, Qahash M, Polimeni G, Leknes KN, Shanaman RH, Prasad HS, Rohrer MD, Hall J. The critical-size supraalveolar peri-implant defect model: characteristics and use. *J Clin Periodontol*. 2006;33:846–54.
- Wilensky A, Polak D, Awawdi S, Halabi A, Shapira L, Houri-Haddad Y. Strain-dependent activation of the mouse immune response is correlated with *Porphyromonas gingivalis*-induced experimental periodontitis. *J Clin Periodontol*. 2009;36:915–21.
- Wirthlin MR, Yeager JE, Hancock EB, Gaugler RW. The healing of gingival wounds in miniature swine. *J Periodontol*. 1980;51:318–27.
- Wyss C, Guggenheim B. Effects of the association of conventional rats with *Actinomyces viscosus* Nyl and *Bacteroides gingivalis* W83. *J Periodontol Res*. 1984;19(6):574–7.
- Wong DM, Tam V, Lam R, Walsh KA, Tatarczuch L, Pagel CN, Reynolds EC, O'Brien-Simpson NM, Mackie EJ, Pike RN. Protease-activated receptor 2 has pivotal roles in cellular mechanisms involved in experimental periodontitis. *Infect Immun*. 2010;78:629–38.
- Yaffe A, Shoshan S. Re-attachment of periodontal ligament by collagen in experimentally-induced alveolar bone dehiscence in dogs. *Arch Oral Biol*. 1987;32:69–73.
- Yamamoto S, Masuda H, Shibukawa Y, Yamada S. Combination of bovine-derived xenografts and enamel matrix derivative in the treatment of intrabony periodontal defects in dogs. *Int J Periodontics Restorative Dent*. 2007;27:471–9.
- Yamamoto T, Tomofuji T, Tamaki N, Ekuni D, Azuma T, Sanbe T. Effects of topical application of lipopolysaccharide and proteases on hepatic injury induced by high-cholesterol diet in rats. *J Periodontol Res*. 2010;45:129–35.
- Yamanaka A, Yasui K, Sonomura T, Uemura M. Development of heterodont dentition in house shrew (*Suncus murinus*). *Eur J Oral Sci*. 2007;115:433–40.
- Yamazaki Y. Effects of destructive periodontitis, induced by diet, on the mechanical properties of the periodontal ligament of the mandibular first molar in golden hamsters. *J Periodontol Res*. 1992;27:149–58.
- Yeo YJ, Jeon DW, Kim CS, Choi SH, Cho KS, Lee YK, Kim CK. Effects of chitosan nonwoven membrane on periodontal healing of surgically created one-wall intrabony defects in beagle dogs. *J Biomed Mater Res B Appl Biomater*. 2005;72:86–93.
- Yu JJ, Ruddy MJ, Wong GC, Sfintescu C, Baker PJ, Smith JB, Evans RT, Gaffen SL. An essential role for IL-17 in preventing pathogen-initiated bone destruction: recruitment of neutrophils to inflamed bone requires IL-17 receptor-dependent signals. *Blood*. 2007;109:3794–802.
- Yu G, Yu Y, Li YN, Shu R. Effect of periodontitis on susceptibility to atrial fibrillation in an animal model. *J Electrocardiol*. 2010;43(4):359–66.
- Zechner W, Tangl S, Tepper G, et al. Influence of platelet-rich plasma on osseous healing of dental implants: a histologic and histomorphometric study in minipigs. *Int J Oral Maxillofac Implants*. 2003;18:15–22.
- Zhang X, Teng YT. Interleukin-10 inhibits gram-negative-microbe-specific human receptor activator of NF- κ B ligand-positive CD4+ Th1-cell-associated alveolar bone loss in vivo. *Infect Immun*. 2006;74:4927–31.
- Zhang MZ, Li CL, Jiang YT, et al. *Porphyromonas gingivalis* infection accelerates intimal thickening in iliac arteries in a balloon-injured rabbit model. *J Periodontol*. 2008;79:1192–9.
- Zhao M, Jin Q, Berry JE, Nociti Jr FH, Giannobile WV, Somerman MJ. Cementoblast delivery for periodontal tissue engineering. *J Periodontol*. 2004;75:154–61.
- Zhou Q, Desta T, Fenton M, Graves DT, Amar S. Cytokine profiling of macrophages exposed to *Porphyromonas gingivalis*, its lipopolysaccharide, or its FimA protein. *Infect Immun*. 2005;73:935–43.
- Zubery Y, Dunstan CR, Story BM, Kesavalu L, Ebersole JL, Holt SC, Boyce BF. Bone resorption caused by three periodontal pathogens in vivo in mice is mediated in part by prostaglandin. *Infect Immun*. 1998;66:4158–62.
- Zucker LM. Hereditary obesity in the rat associated with hyperlipemia. *Ann N Y Acad Sci*. 1965;131:447–58.

Index

A

Absolute risk difference, 193
Absorbent filter paper strip, 377, 389
Accuracy, 20, 34, 39, 155, 159, 161, 166, 167, 172, 178, 228, 316, 324, 326, 328, 330, 344, 346, 353, 365, 390, 391, 393, 423–440, 460, 488, 493, 507, 560, 581, 583, 586, 613–615, 617, 620, 628, 634
Accutek probe, 162, 321
Acoustic microscopy, 342–346
Actinobacillus actinomycetemcomitans, 352, 361, 540, 594, 604, 694, 697
Actinomyces israelii, 697, 698
Actinomyces naeslundii, 179, 348, 697, 698, 700, 704, 705
Actinomyces odontolyticus, 697
Actinomyces viscosus, 679, 697, 698, 704
Acute/chronic defect model, 658, 671, 673
Acute defect model, 671
Adaptive bias-coin designs, 93
ADA seal of acceptance, 450, 473
Agency for Healthcare Research and Quality (AHRQ), 235, 290, 292
 checklist, 199, 203, 235, 290, 292, 293, 433
Aggressive periodontitis, 73, 115, 245, 267, 297, 311, 348, 371, 375, 400, 401, 477, 484, 488, 490
AHRQ. *See* Agency for Healthcare Research and Quality (AHRQ)
Alabama probe, 162
Allocation concealment, 69, 73, 76, 78, 92, 95–97, 124, 128–130, 132, 136, 145, 155, 156, 207, 209, 211–214, 237, 743
Allocation ratio, 109, 124
A Measurement Tool to Assess Systematic Reviews (AMSTAR), 232, 237, 277, 745
AMSTAR. *See* A Measurement Tool to Assess Systematic Reviews (AMSTAR)
Antigingivitis agents, 441, 449, 450
Antimicrobial agents, 90, 185, 441, 560
Antiplaque agents, 441, 443, 461, 543
Appraisal of Guidelines, Research and Evaluation (AGREE)
 collaboration, 287
 instrument, 287, 289, 294
A priori power analysis, 258
ARCPOH. *See* Australian Research Center for Population Oral Health (ARCPOH)

ARRIVE Guidelines, 741–744
Ascertainment bias, 95, 152, 154
Assessment of Multiple Systematic Reviews tool, 232
Atomistic fallacy, 299
Attrition bias, 152–153, 202, 212, 215, 238
Australian Research Center for Population Oral Health (ARCPOH), 78
Authorship, 12, 13, 21–25, 90, 273
Available case analysis, 114

B

Bacterial culture, 354, 356, 361, 365
Bacterial vitality within the dental plaque, 560–562
Belmont Report, 11
Bias
 information, 34
 in sampling, 579
 selection, 146, 212, 233, 439
Biased-coin designs, 93
Bite-wings, 326
Bizzini scale, 208
Blinding, 5, 19, 26, 76, 91, 95, 97–106, 124, 128–132, 151, 154, 155, 209–213, 215, 217, 218, 238, 451, 741, 743, 744, 746
Burst
 model, 302, 619
 theories, 303, 621

C

CADIA. *See* Computer-assisted densitometric image analysis (CADIA)
CAL. *See* Clinical attachment level (CAL)
Calculus index, 46, 168, 517, 540, 541, 563, 607
Calculus rating, 53
Calculus surface index (CSI), 563
Calibration, 133, 159, 161, 163–167, 172, 174–178, 316, 324, 390, 391, 393, 513, 521, 522, 530, 558, 590–593, 621, 623
Capillary tubing, 389
Care effect, 459
Carry-across effect, 74–76
Carry-over effect, 60, 73–75
Case-cohort design, 576

- Case-control study, 3, 5, 34, 36, 42–56, 60, 65–67, 82–84, 146, 147, 200, 202, 204, 206, 286, 428, 482, 488, 490–492, 494–498, 501, 575, 576, 589, 594–596, 598, 601, 602, 636, 650, 677
- Case definition, 35, 535, 589–618, 633, 635
- Case-only design, 202, 482, 489
- Case report, 12, 13, 15, 34–36, 126, 166, 202
- Case-series report, 34–37
- CASP. *See* Critical Appraisal Skills Programme (CASP)
- CASP appraisal tool, 433
- CBCT. *See* Cone-beam Computed tomography (CBCT)
- CDC/AAP, 613, 615
- CDC Periodontal Disease Surveillance Project, 627–628
- Cetylpyridinium, 469, 472
- Chair-side dipstick test for MMP-8, 402, 403
- Chalmers scale, 209, 211
- Channeling bias, 146
- CHD. *See* Coronary heart disease (CHD)
- Checkerboard hybridization, 346, 358–361, 601, 602
- Chlorhexidine, 104, 115, 186, 220, 222, 223, 229, 259, 267, 443, 445, 446, 452, 461, 469–472, 689
- Chronic periodontitis, 2, 44, 58, 59, 61, 115, 164, 266, 267, 291, 311, 348, 349, 351, 352, 375, 380, 394, 400, 401, 470, 473, 477, 484, 488, 490, 539, 611, 648, 726
- Chronology bias, 146, 147
- CIOMS. *See* Council for International Organisations of Medical Sciences (CIOMS)
- Citation Bias, 153–154
- Clinical attachment level (CAL), 44–49, 51, 55, 105, 112, 115, 121, 149–151, 155, 156, 220–226, 246, 261, 262, 265–267, 283, 307, 308, 310, 315, 319, 321, 324, 325, 374, 375, 379–380, 396–400, 442, 527, 529, 530, 532, 537–539, 579, 582–584, 586, 587, 589–595, 597–599, 602–606, 608–618, 694
- Clinical modelling, 657
- Clinical practice guidelines, 248, 271–294, 438–440
- Clinical question, 1–2, 4, 26, 191, 195–196, 199, 261, 288, 438
- Clinical Trial Registration, 115–116, 128
- Cluster sampling, 177, 578
- CMA. *See* Comprehensive Meta-Analysis (CMA)
- Cochrane Collaboration, 200, 206–210, 212, 218, 247, 433, 435–436, 744
- Cochrane Collaboration Guidelines, 82
- Cochrane Handbook, 114, 218, 435, 436
- COGS. *See* Conference on Guideline Standardization (COGS)
- Cohen's kappa, 170–172
- Cohort studies, 2, 3, 5, 33, 42, 43, 52, 54–62, 64–66, 70, 83, 197, 202, 204, 207, 286, 482, 483, 493–496, 501, 575–576, 586, 590, 635
- Collagenase-2, 401
- Community Periodontal Index (CPI), 379, 380, 527, 566, 586, 591, 630, 631
- Community Periodontal Index of Treatment Needs (CPITN), 44–46, 50, 51, 55, 319, 379, 380, 528, 564–566, 580, 582, 593, 602, 603, 629, 631, 632
- Complex disease, 477–482, 484, 488, 490, 492, 505
- Compliance, 97, 109, 112, 121, 131, 132, 185–189, 201, 227, 266, 637, 653
- Composite endpoint, 149
- Comprehensive Meta-Analysis (CMA), 230, 231, 249
- Comprehensive smoking index (CSI), 648–649, 660
- Compromised power analysis, 258
- Computer-assisted densitometric image analysis (CADIA), 326, 328, 677
- Cone-beam computed tomography (CBCT), 330–331, 335, 337, 338
- Conference on Guideline Standardization (COGS), 277, 286
- Confidence interval (CI), 34, 42, 44–51, 54, 62, 63, 67, 68, 71, 75, 82, 83, 98, 105, 109, 124, 128, 131, 133–136, 145, 161, 165, 172, 176, 193, 195, 226–229, 233, 234, 245, 262, 264, 284, 299, 301, 302, 304, 310, 311, 340, 341, 352, 429, 431, 432, 437, 438, 496, 583, 589, 603, 611, 612, 646, 647, 743, 745, 746
- Conflicts of interest, 12, 13, 19–20, 122, 154–155, 204, 238, 241, 273, 275–276, 288, 505
- Confounder, 60, 62, 65, 66, 146, 202, 203, 206, 207, 262, 298, 303, 492, 494, 496, 646–649, 651, 652
- Confounding, 33, 34, 54, 57, 64, 65, 70, 71, 75, 82, 91, 92, 146, 151, 203, 204, 281, 285, 366, 439, 482, 483, 487, 492, 494, 495, 498, 500, 577, 645–653, 658, 746
- Consecutive sampling, 40, 82, 431, 579
- Consistency, 10, 102, 159, 233, 234, 271, 272, 283, 286, 319, 321, 347, 370, 423, 430, 486, 513, 534, 543, 616, 623, 698
- Consolidated Standards of Reporting Trials (CONSORT) guidelines, 76, 103, 123–133
- statement, 69, 103, 113, 123–129, 131, 133, 154, 208, 741
- Consolidated standards of reporting trials (CONSORT) statement, 69, 103, 113, 124–129, 131, 133, 154, 208, 741
- Construct validity, 40
- Content validity, 40, 237
- 2 x 2 Contingency table, 429, 652
- Continuous variable, 46, 66, 107, 168, 193, 195, 230, 496, 539, 648
- Convenience sampling, 80, 578
- Coronary heart disease (CHD), 43, 54, 55, 61–62, 479, 607, 646, 647
- Cost-benefit analysis, 63
- Cost-effectiveness analysis, 61, 274
- Council for International Organisations of Medical Sciences (CIOMS), 10, 118
- Covariance components models, 298, 300, 301
- CPI. *See* Community Periodontal Index (CPI)
- CPITN. *See* Community Periodontal Index of Treatment Needs (CPITN)
- CPITN Teeth, 580, 582
- Criterion validity, 40
- Critical appraisal, 2–4, 69, 82, 191, 203, 243, 272, 493
- Critical Appraisal Skills Programme (CASP), 218, 433, 434

Critical reading, 3
 Critical-size defect models, 657
 Critical-size supra-alveolar peri-implant defect model, 663, 670
 Critical size supra-alveolar periodontal defect model, 173, 663
 Cross-over design, 72–75, 112, 202, 445
 Cross-sectional, 3, 33–35, 42–43, 60, 64–67, 82, 84, 200, 202, 204, 207, 300–302, 316, 373–375, 378–390, 396–400, 428, 493–497, 501, 575–577, 594–597, 599, 600, 602, 603, 635, 636
 CSI. *See* Calculus surface index (CSI)
 Cumulative meta-analysis, 244–247, 249
 Cytofluorography, 356

D

Darkfield microscopy, 352
 Data combination, 194, 197
 Data pooling, 194
 Data safety monitoring boards (DSMBs), 19, 26, 27, 90
 Decision matrix, 429
 Declaration of Helsinki (DoH) of the World Medical Association (WMA), 10, 449
 Delphi list, 208, 209, 211
 DentoAnalyzer, 402, 404
 Design biases, 147–151
 Detection bias, 151, 202, 212, 215, 437
 Diagnostic odds ratios (DORs), 365, 433, 437
 Diagnostic studies, 71, 146, 423, 427, 430, 434–438
 Diagnostic test, 1, 197, 207, 279, 285, 364, 371–372, 388, 423, 424, 426–438, 440, 618, 636, 637
 Diagnostic trials, 31
 Diamines, 466
 Diamond Probe®/Peri2000® System, 469
 Dichotomous variable, 108, 145, 194, 195, 539
 Digital subtraction radiography (DSR), 326–328, 331, 335
 Digit preference error, 161
 DNA-DNA hybridization methods, 90, 346, 357–363
 Dogs, 323, 339, 655, 656, 659, 676–692, 704, 730, 737
 DORs. *See* Diagnostic odds ratios (DORs)
 DSMBs. *See* Data Safety Monitoring Boards (DSMBs)
 DSR. *See* Digital subtraction radiography (DSR)
 Dutch Cochrane checklist, 204

E

Eastman interdental bleeding index (EIBI), 515, 519
 EasyMA, 230, 231
 Ecological fallacy, 299
 Economic and decision analyses, 63
 Effect modification, 645–653
 Effect size, 5, 81, 103, 109, 122, 125, 130, 155, 156, 160, 215, 216, 239, 246, 249, 258–260, 437, 461, 484, 489, 490, 493, 502
 EGCS. *See* Ethical Guidelines for Clinical Studies (EGCS)

Egger's linear regression approach, 240
 EIBI. *See* Eastman interdental bleeding index (EIBI)
 Electronic searching, 199–200, 243
 ELISA. *See* Enzyme-linked immunosorbent assay (ELISA)
 Enhancing the Quality and Transparency of Health Research Network (EQUATOR), 127
 Enzyme-linked immunosorbent assay (ELISA), 355, 356, 384–386, 403, 404, 677, 725, 729
 EP table, 280, 282
 EQUATOR. *See* Enhancing the Quality and Transparency of Health Research Network (EQUATOR)
 Equivalence trial, 123, 125, 128, 134–137, 453
 Essential oils, 446, 469, 471
 Ethical Guidelines for Clinical Studies (EGCS), 9
 Ethnic heterogeneity, 487–488
 Evalusite® Test, 356
 Evidence-based dentistry, 3, 6, 199
 Evidence based medicine, 1–6, 195, 438
 Evidence based periodontology, 1–6
 Experimental gingivitis model, 448
 Experimentally-produced furcation defects, 679
 Experimental periodontitis, 339, 671–674, 676–692, 694, 696–704, 706–712, 714, 716, 721, 727, 730, 731, 733
 Extended follow-up design, 72
 External validity, 57, 66, 79, 82, 113, 125, 131, 201, 202, 238, 428, 430, 438, 439, 496, 602, 650

F

Face validity, 40
 Factorial design, 72
 Family-based designs, 482, 483
 Fenestration defect, 331, 658, 673, 674, 688, 689, 703, 727
 Ferret, 34, 655, 656, 730
 Fixed effects, 194, 230, 299–302, 311
 Fixed sites full-mouth, 580–581
 Fleiss' kappa, 171
 Florida probe system, 162, 320
 Flow cytometry, 356
 Flow diagrams, 66, 69, 113, 114, 123, 126, 130, 154, 232, 233, 430–432, 434, 495
 Fluorescence in situ hybridization (FISH), 357–358
 Follow-up adherence, 185, 187
 Forest plot, 194, 227, 228, 233
 Four-days plaque re-growth model, 444
 Fractal analysis, 328–329
 Full-mouth model, 447
 Funding bias, 238
 Funnel plot analysis, 122, 194, 239
 Furcation, 36, 170, 260, 318, 331, 332, 527, 535–537, 539, 566, 567, 592, 604, 635, 637, 657, 658, 674, 675, 679, 680, 692, 694, 695, 703, 708–713, 729, 731

G

Gas chromatography, 467, 468, 470–472, 475
 GBI. *See* Gingival bleeding index (GBI)
 GCF. *See* Gingival crevicular fluid (GCF)
 Genetic biomarkers, 377, 378
 Genetic-environment (g x e) interactions, 481, 482, 489, 503
 Gingival bleeding index (GBI), 51, 448, 513–515, 631
 Gingival bleeding time index, 514–515
 Gingival crevicular fluid (GCF), 91, 160, 316, 347, 372, 378, 383, 388–404, 448, 450, 507, 511, 595–597, 600, 603, 610, 611, 677
 Gingival index (GI), 112, 167, 172, 261, 380, 400, 510–516, 519, 524, 603, 608, 632, 694, 731
 Gingival washing methods, 388
 Gold standard, 2, 70, 82, 89, 98, 138, 159, 167, 196, 200, 202, 203, 207, 259, 330, 354, 385, 429, 437, 468, 492, 532, 537, 565, 613, 614, 618, 628–632, 635
 GRADE. *See* Grading of recommendations assessment, development and evaluation approach (GRADE)
 Gradepro computer software, 283
 Grading of recommendations assessment, development and evaluation approach (GRADE), 278–285, 438, 440
 Grey literature, 2, 199, 200, 237, 241, 248
 Grips statement, 497–503
 Group-randomized trials (GRTs), 137–138
 GRTs. *See* Group-randomized trials (GRTs)
 Guideline grading system, 278
 Guidelines for the WHO guidelines (GWG), 275, 277, 278, 289
 Guidelines on good publication and the code of conduct for editors of biomedical journals of the committee on publication ethics (COPE), 10, 11
 GWG. *See* Guidelines for the WHO guidelines (GWG)

H

Hamster, 655–657, 714–715, 733
 Hawthorne effect, 443, 459–462
 Hawthorne stabilization, 461
 Haynes's 5 S's, 200–201
 Healthy people 2010 (Hp 2010), 613, 615, 623, 626, 627
 Hedonic measurement, 468
 Heterogeneity, 75, 148, 192, 194, 195, 197, 209, 228–230, 236, 238, 242, 246, 249, 265, 364, 435, 436, 438, 439, 487–488, 490, 590, 652, 745–747
 Hispanic hanes, 622
 Homogeneity, 192, 194, 238, 382, 747
 Human Genome Project, 478–479

I

ICC. *See* Intraclass correlation coefficient (ICC)
 ICMJE. *See* International committee of medical journal editors (ICMJE)
 Immunofluorescence, 346, 356

Imprecision, 5, 6, 66, 90, 125, 131, 145, 279, 281–283, 285, 496, 743
 Individual patient data (IPD) meta-analysis, 192, 244
 Information biases, 151–152
 Informed consent, 9–13, 15–19, 27, 80, 116–122, 157, 166, 186, 187, 241
 Inoculation of microorganisms, 657
 Institutional review boards (IRB), 11, 13–15, 17, 18, 26, 90, 115–117, 121
 Intention-to-treat analysis (ITT), 36, 112–114, 126, 129, 130, 132, 135, 154, 220–226, 238
 Interdental cratering, 705
 Inter-examiner reliability, 159, 161, 165, 167, 168, 177, 517
 Internal validity, 3, 4, 34, 82, 113, 120, 152–154, 201–204, 283
 International committee of medical journal editors (ICMJE), 10–13, 21–25, 115, 116, 123, 128, 153, 154
 Interproximal periodontal defect (IPD), 681, 687
 Intraclass correlation coefficient (ICC), 138, 168–170, 174, 224, 586, 623, 624, 626
 Intra-examiner reliability, 159, 161, 165, 167, 168, 170, 172, 177, 517
 Intraosseous defects, 674, 681
 Investigator integrity, 20–21
 IPD. *See* Interproximal periodontal defect (IPD)
 IRB. *See* Institutional review boards (IRB)
 ITT. *See* Intention-to-treat analysis (ITT)

J

Jadad scale, 208, 209, 211

K

Kappa scores, 170–172, 626

L

Ligature-induced defects, 657
 Likelihood ratio, 426, 428–429, 437
 Linear model, 298, 339
 Linear theories, 302
 Line of no effect, 193, 228, 229

M

Maastricht scale, 208
 Magnetic resonance imaging (MRI), 332, 336–338
 Maher Maastricht-Amsterdam list (MAL), 208, 209
 MAL. *See* Maher Maastricht-Amsterdam list (MAL)
 Mantel-Haenszel procedure, 651
 Manual of operations, 115
 Masking, 72, 76, 91, 96, 97, 102, 104, 130, 148, 155, 156, 317, 326, 449, 469, 553
 Matching, 5, 65, 67, 95, 100, 203, 361, 492, 495, 650, 743
 Mathematical coupling (MC), 305–307, 310
 Matrix metalloproteinases (MMPs), 371, 378, 388, 393, 401, 403, 484, 702, 712

Mendelian patterns of inheritance, 477, 482
 Meta-analysis, 54, 55, 61, 75, 96, 102, 145, 149, 155, 156, 192–195, 197, 213, 216, 227–235, 239, 241–247, 249, 262, 277, 286, 289, 300, 365, 434–439, 484, 537, 619, 745, 746
 Meta-analyst, 229, 239
 Meta-epidemiology, 145
 MetaAnalysis, 230, 231
 Meta-narrative reviews, 243
 Metawin, 230, 231
 Microbial biomarkers, 377, 378
 Microscopic methods, 353
 Minimization, 97, 214, 216, 274
 MIQE, 367
 Mix 1.5, 230, 231
 Mixed-effects models, 298
 Mixed randomization, 93, 94
 MLwin software, 300
 Mobility index, 507, 564
 Moose statement, 232
 MRI. *See* Magnetic resonance imaging (MRI)
 Multi-level linear models, 298
 Multilevel modelling, 297–312, 620
 Multiplex PCR, 346, 364
 Multistage sampling, 578–579
 Myperiod[®] pst[®], 387
 Myperiopath[™], 365

N
 Narrative review, 191, 194, 242, 243
 National guideline clearinghouse (NGC), 199, 271, 291
 National Health and Nutrition Examination Survey (NHANES)
 Nhanes I, 622, 626, 627
 Nhanes II, 621, 627
 Nhanes III, 47, 49, 51, 157, 165, 527, 535–537, 581, 586–588, 590, 613–615, 621–627, 648
 Navy periodontal disease index (NPDI), 527, 534
 Navy plaque index, 534, 544–546, 550
 Near infrared spectroscopy (NIR), 338–341
 Negative predictive value, 424, 428, 514, 532, 565, 614, 632, 637
 Nested case-control study, 5, 60, 482
 Network meta-analyses, 243–244
 Newcastle-Ottawa quality assessment scale, 82, 83, 205
 Newcastle-Ottawa scale (NOS), 204
 New method of plaque scoring (NMPS), 546–547
 NGC. *See* National guideline clearinghouse (NGC)
 NHANES. *See* National Health and Nutrition Examination Survey (NHANES)
 NMPS. *See* New method of plaque scoring (NMPS)
 NNT. *See* Number needed to treat (NNT)
 No-brushing studies, 443
 Nomograms, 110, 263, 264, 429
 Non-adherence, 112, 114, 185, 186, 188, 189
 Non human primates, 656, 675, 677, 684
 Non-inferiority trials, 76, 134–135

Non-randomized controlled trial, 3, 5, 69
 Non-randomized trial, 36, 64–70
 Nucleic acid probe method, 357
 Null hypothesis, 33, 78, 89–90, 108, 135–137, 236, 257, 258, 306, 493, 589
 Null model, 298, 299, 301, 307
 Number needed to treat (NNT), 112, 194, 262–266
 Nuremberg code (directives for human experimentation), 9, 10, 116

O

Observational Clinical Research, 34–64
 OCT. *See* Optical coherence tomography (OCT)
 Odds ratio (OR), 34, 50, 51, 54, 60, 68, 81, 82, 96, 102, 145, 179, 192, 194, 202, 206, 207, 226–229, 245, 304, 305, 368, 426, 428, 433, 437, 484, 488, 489, 493, 502, 519, 576, 583, 590, 591, 603, 611, 612, 616, 635, 636, 650, 651, 747
 OFNASET. *See* Oral fluidic nanosensor test (OFNASET)
 Oligonucleotide microarray technology, 361–363
 One-wall defect, 681
 Ontology, 381–383
 Optical coherence tomography (OCT), 331–335
 Optimism bias, 154
 Option-3 scheme, 178
 OQAQ. *See* Overview quality assessment questionnaire (OQAQ)
 Oral calculus index, 563
 Oral fluidic nanosensor test (OFNASET), 383, 384
 Oral rating index (ORI), 516–517
 Organoleptic measurement, 467, 468
 Overarching bias, 154–155
 Overview quality assessment questionnaire (OQAQ), 235, 237

P

Papillary bleeding index (PBI), 380, 400, 511, 631
 Parallel group design, 3, 71
 Parocheck[®], 362
 Partial recording protocol, 579–589
 PBI. *See* Papillary bleeding index (PBI)
 PCR method, 363–370
 PDI. *See* Periodontal disease index (PDI)
 PDR. *See* Periodontal disease rate (PDR)
 PEDro scale, 208, 209, 211
 Percentage agreement, 168, 175, 514
 Performance bias, 146, 152, 202, 212, 215
 Peri-implantitis, 150, 339, 341, 342, 401, 402, 672, 690–694, 747
 Periodontal Disease Index (PDI), 51, 527, 529–530, 593
 Periodontal disease rate (PDR), 527, 530–531
 Periodontal index, 261, 526, 528, 529, 564, 566
 Periodontal Inflamed Surface Area, 537–541
 Periodontal Pocket Bleeding Index (PPBI), 55, 515
 Periodontal screening and recording index (PSR), 528, 566–567
 Periodontitis severity index (PSI), 519, 531

- Periotemp™, 315
 Per-protocol analysis, 114, 135
 Phase contrast microscopy, 353
 Phenyl compounds, 466
 Phylochip®, 362
 PICO acronym, 1, 196
 PICOTT, 197
 Placebo(s)
 active, 26
 inactive, 99
 Plaque area measurements, 549–558
 Plaque control record, 544
 Plaque index, 44, 48, 49, 51, 55, 149, 168, 172, 174, 226, 229, 301, 302, 304, 305, 310, 312, 380, 400, 444, 449, 450, 534, 541–543, 545, 550, 583, 603, 611, 612, 694
 Plaque reducing/inhibitory agents, 441
 PMA index, 509, 510
 Point estimate, 96, 134, 228, 265, 589
 Point-of-care diagnostics, 383–388
 Polygenic inheritance, 482
 Population-based designs, 482, 483
Porphyromonas gingivalis, 347, 348, 354, 357, 361, 377, 388, 466, 540, 594, 604, 694, 697–699, 702, 716–718, 720, 740, 741
 Positive predictive value (PPV), 428, 532, 632, 637
 Post-hoc power analysis, 258
 Post-test probability of disease, 429
 Power analysis, 258, 259
 Power of a study, 258
 PPBI. *See* Periodontal Pocket Bleeding Index (PPBI)
 PPV. *See* Positive predictive value (PPV)
 Precision, 34, 66, 81–83, 95, 107, 109, 111, 125, 131, 161, 177, 193, 194, 200, 227, 228, 234, 239, 264, 320, 321, 343–344, 363, 426, 496, 504, 530, 551, 583, 634, 653, 743
 Predictive validity, 40, 633
 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), 232, 243
 Prevention trials, 31
 PRISMA. *See* Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)
 Probe tine diameter, 161
 Probing depth, 2, 51, 55, 80, 115, 149, 150, 160, 162, 164, 165, 168, 170, 176, 179, 226, 261, 265, 266, 268, 283, 297, 301–305, 308, 316, 318, 319, 321–325, 329, 351, 354, 375, 380, 391, 400, 442, 467, 473, 509, 511, 528, 535–537, 567, 583, 586, 589–593, 603, 611, 614, 627, 633, 635, 652, 694, 695
 Prognostic studies, 437
 Propensity score, 203, 651
 Proportions, 9, 32, 43, 80, 82, 98, 111–113, 115, 128, 130, 150, 168–171, 176, 177, 200, 202, 211, 213, 215, 216, 262, 264, 349–352, 354, 368, 369, 390, 428, 430, 439, 444, 466, 479, 490, 519, 520, 522, 532, 565, 581, 605, 606, 609, 619, 636, 680, 707, 729
 Prospective meta-analysis, 244
 Proteomic biomarkers, 376–378, 382, 384
 Protocol adherence, 185
 Protocol effect, 459
 Pseudohalutosis, 465
 PSI. *See* Periodontitis severity index (PSI)
 PSR. *See* Periodontal screening and recording index (PSR)
 Publication bias, 115, 116, 122, 153, 194, 233, 239–241, 247–249, 279, 281, 282, 292, 293, 438, 439, 493, 745
- Q**
 Q-test, 194
 QUADAS. *See* Quality assessment of diagnostic accuracy studies (QUADAS)
 Quality assessment of diagnostic accuracy studies (QUADAS), 433–435
 Quality of life trials, 31
 Quality of reporting of meta-analyses (QUOROM), 232, 437, 438
 statement, 232–234, 438
 Quality of the study, 194, 197
 Quality score, 62, 128, 132, 194, 289
 Quantitative study design stages, 33
 Quasirandomized, 194
 QUOROM. *See* Quality of reporting of meta-analyses (QUOROM)
 Quota sampling, 80
- R**
 Rabbit, 655, 727–728, 733
 Ramfjord teeth, 527, 529, 530, 534, 580, 582, 584, 586, 591
 Random-coefficient regression models, 298
 Random coefficients model, 300, 302, 303, 620
 Random effects, 138, 169, 299–302, 307, 311
 Random effects model, 195, 228, 230, 238, 246, 298
 Random half-mouth, 580–582, 586, 587
 Random intercept, 300–302, 309
 Randomization, 3, 17, 22, 70–74, 76, 78, 91–97, 105, 113, 117, 125, 128–132, 136–138, 146, 148, 152, 154, 186–187, 203, 213, 214, 216–218, 427, 450, 451, 483, 497, 519, 649, 741–744, 746
 Randomized controlled trial (RCT), 2, 3, 22, 33, 67, 70, 82, 89, 96, 100, 106, 112–115, 121–123, 127–133, 137, 149, 189, 200, 202, 207–213, 228, 232–234, 245, 262, 280, 285, 286, 307, 500
 Random permuted blocks, 78, 92, 94, 95
 Random subsets of teeth (or sites), 581
 Rank correlation test, 239–240
 Rat, 77, 333, 337, 580, 656–659, 694–714, 719, 727, 731–733, 737, 738
 Ratios, 50, 51, 58, 60, 63, 81, 82, 102, 109, 112, 122, 145, 174, 177, 193, 194, 206, 227–229, 233, 240, 262, 265, 268, 286, 299, 305, 326, 335, 339, 371, 424, 428–429, 437, 467, 508, 519, 531, 576, 577, 597, 603, 632, 635, 636, 646, 663, 666, 670, 745
 RCT. *See* Randomized controlled trial (RCT)
 Realist and metareview evidence synthesis: evolving standards (RAMESES), 243

- Realist reviews, 242
 Real-time PCR, 349, 353, 364–367
 Receiver operating curves (ROC), 386, 429, 435–437, 633
 Recruitment, 65, 75, 80, 89, 90, 102, 108, 111, 112, 116, 125, 130, 147, 187, 192, 431, 494, 504, 540, 718
 Regimen adherence, 185
 Regression modeling, 58, 59, 433, 594, 604, 620, 633, 648, 650
 Regression to the mean (RTM), 60, 303, 305–307, 621
 Relative attachment level (RAL), 163, 319, 321, 324, 325, 608, 611, 620
 Relative DORs, 433
 Relative risk, 34, 50, 51, 54, 57, 58, 60, 62, 66, 67, 81, 112, 122, 193, 194, 202, 226–228, 261–265, 282, 428, 479, 480, 496, 576, 590, 611, 616, 635, 636, 652
 Reliability, 3, 114, 152, 154, 159–179, 204, 237, 244, 260, 271, 325, 390, 393, 423, 426, 515–517, 532, 546, 548, 555, 621, 623, 625–627
 Replacement randomisation, 95
 Response bias, 37, 100
 Restricted randomization, 92–94
 Restriction, 94, 124, 130, 202, 221, 367–368, 443, 649, 745
 Retention index system, 507, 563
 Review manager (Revman), 213, 218, 230, 231
 Risk ratio, 58, 81, 194, 227, 233, 240, 265, 284, 576
 ROC. *See* Receiver operating curves (ROC)
- S**
 Saliva, 160, 353, 355, 368, 370, 372–388, 392, 396, 397, 399, 403, 404, 444, 445, 466, 474, 529, 610
 Salivary bacterial counts, 445
 Sample size, 12, 33, 90, 148, 154, 211, 258, 348, 435, 451, 483, 535, 577, 649, 742
 Sample size calculation, 76, 80, 106–112, 137, 148, 211, 258, 742, 744
 Sampling
 snowball, 80, 578–579
 stratified, 80
 weights, 579, 646
 SAS, 300, 623, 625, 626
 SBI. *See* Sulcus bleeding index (SBI)
 Scientific misconduct, 20
 Screening models, 657, 658
 Screening test, 426, 437
 Selection bias, 34, 36, 65, 69, 78, 80, 91, 95, 96, 138, 145, 151, 153, 154, 202, 203, 206, 207, 212, 214, 235, 240, 244, 439, 499, 590
 Sensitivity, 61, 63, 65, 66, 73, 136, 193, 195, 197, 200, 233, 234, 241, 249, 326, 330, 335, 337, 346, 354–356, 364, 365, 368, 370, 371, 385, 386, 401, 402, 424, 427–431, 433, 436–438, 468, 495, 496, 515, 516, 520, 551, 560, 566, 567, 581, 585, 587, 606, 614, 615, 620, 628, 632, 633, 636, 637, 737
 Sequence generation, 76, 95, 124, 128, 130, 209, 212, 214, 218, 744
 Simple randomisation, 92, 95
 Simple random sampling, 80, 577
 Simplified oral hygiene index, 540–541
 Smoking, 3, 34, 105, 147, 160, 219, 262, 300, 448, 470, 477, 540, 577, 646
 SoFs table, 280, 283
 Specificity, 197, 199, 249, 355–357, 364, 365, 370, 384, 386, 401, 402, 424, 427–431, 433, 436–438, 514, 532, 548, 564, 565, 606, 614, 628, 632, 633, 635–637
 Spill-over effects, 75
 Split-mouth/cross-over, 74–78
 Split-mouth design, 3, 74–78, 148, 447, 675
 Split-plot design, 74
 Standard error of measurement, 168–169, 325
 Standardization, 148, 161, 166–167, 172, 173, 277, 286, 327, 328, 348, 389, 451, 520, 623, 625, 658, 712
 Standardized mean difference, 193, 195
 Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA), 126
 Standards for Reporting of Diagnostic Accuracy (STARD), 430–432, 497, 503
 STARD. *See* Standards for Reporting of Diagnostic Accuracy (STARD)
 Stata, 100, 300, 578
 Statistical significance, 75, 108, 134, 135, 149, 153, 167, 171, 228, 229, 241, 242, 246, 257–260, 267, 303, 471, 472, 479, 596, 598, 607, 655, 657
 Stratification, 94, 95, 97, 130, 146, 203, 482, 483, 493–495, 498, 500, 650, 652, 653
 Stratified randomization, 94–95, 483, 649
 Stratified random sampling, 577–578
 STREGA statement, 493–497
 Strengthening the reporting of genetic association (STREGA) studies, 493–503
 Strengthening the reporting of observational studies in epidemiology (STROBE), 63–64, 66, 154, 493–497
 statement, 63–64, 493–497
 Strength of evidence, 4, 234
Streptococcus mutans, 361, 694, 697, 700, 704
 STROBE. *See* Strengthening the reporting of observational studies in epidemiology (STROBE)
 Stroke, 61, 62, 172, 352, 484, 514, 515, 646, 647, 746
 Subclassification, 650, 651
 Subgingival temperature, 315–316
 Subgroup analysis, 130, 195, 197, 439
 Submerged wound closure, 659
 Sulcus bleeding index (SBI), 151, 400, 508, 510–511, 695
 Sulphide monitors, 468, 475
 Summary estimate, 195, 433, 436, 651, 652
 SUMSearch, 199
 Superiority trial, 76, 133–137, 453
 Surrogate endpoints, 91, 148–150, 241–242, 261, 262
 Survey, 3, 15, 33, 97, 149, 164, 369, 508, 575–637, 743
 Susceptibility biases, 145–147
 Systematic random sampling, 577
 Systematic review, 2, 5, 65, 69, 82, 121, 122, 129, 149, 155, 162, 191–249, 274, 275, 277–280, 283, 286, 289, 291–293, 324, 433–439, 470, 473, 590, 628, 744–745
 Systematic sampling, 82

T

- TACT. *See* Tuned aperture computed tomography (TACT)
- Terminal restriction fragment length polymorphism (T-RFLP), 367–368
- The 8-hour substantivity studies, 444
- The 24-h plaque re-growth model, 444
- Therapeutic studies, 75, 204
- The three weeks studies, 470, 474, 475, 671, 677, 691, 694, 706
- Three-arm study, 72
- Three-walls defect, 675, 681
- Toronto Probe, 162, 321
- Transfer bias, 146, 152
- Transparent reporting of evaluations with nonrandomized designs (TREND), 69
- Treatment effect, 26, 64, 69, 70, 73–76, 90, 96, 103, 107, 112, 122, 128, 135–137, 153–155, 167, 185, 186, 207, 227, 228, 239, 245, 249, 262, 305, 307, 459, 461, 497, 500, 512, 635, 651, 745–747
- Treatment-received analysis, 114
- Treatment trials, 18, 31
- TREND. *See* Transparent reporting of evaluations with nonrandomized designs (TREND)
- T-RFLP. *See* Terminal restriction fragment length polymorphism (T-RFLP)
- Triclosan, 310, 312, 448, 469, 472
- TRIM and FILL method, 241
- TRIPdatabase, 199
- Trohoc fallancy, 576
- True endpoints, 148–150, 241, 261, 262
- Tuned aperture computed tomography (TACT), 335–336
- Twarm design,
- Two-walls defect, 675, 681
- Type I error, 107–109, 195, 240, 249, 257–259, 298, 502, 589
- Type II error, 107, 108, 195, 257–259, 298

U

- Ultrasonographic (US) probe, 322, 323, 344
- Ultrasound imaging, 342, 343, 346
- Universal Declaration on Bioethics and Human Rights, 10

V

- Validity, 2, 34, 97, 153, 159, 200, 271, 356, 430, 496, 517, 576, 650, 681
- Van Tulder scale, 208, 209
- Variable, 33, 95, 145, 161, 191, 257, 297, 316, 443, 480, 514, 575, 645, 657
- Variance, 75, 108, 160, 166, 169, 171, 172, 175, 177, 178, 194, 195, 228, 239, 240, 247, 249, 288, 298–302, 305, 307, 309–312, 324, 325, 484, 490, 491, 579, 589, 616, 620, 651, 681, 704, 740
- Variation, 15, 60, 78, 90, 97, 129, 145, 159–160, 163, 165, 169, 173, 177, 178, 228, 230, 236, 271, 279, 286, 287, 298–301, 303, 304, 310, 322, 324, 326, 332, 349, 352, 373, 378, 383, 390, 392, 393, 432–433, 436, 439, 477–479, 481, 484, 487, 488, 490, 497, 499, 500, 503, 529, 532, 577, 589, 605, 617, 620, 621, 635, 647, 649, 652, 653, 655, 673, 727, 746
- Vital fluorescence, 446, 560–562
- Volatile sulphur compounds, 465–472, 723

W

- WAME. *See* World Association of Medical Editors (WAME)
- Washout period, 60, 72, 74, 471
- WEasyMA, 230–231
- Weighted kappa, 168, 170–172
- Weighted mean difference (WMD), 195, 228–229
- Weighting, 82, 194, 195, 204
- WHO. *See* World Health Organisation (WHO)
- Withdrawals, 17, 72, 112, 115, 117, 119, 128, 129, 132, 211, 217
- WMD. *See* Weighted mean difference (WMD)
- World Association of Medical Editors (WAME), 10, 11
- World Health Organisation (WHO), 10, 44–47, 50, 51, 116, 153, 161, 275–278, 287, 289, 319, 425, 527, 564–567, 592

Y

- Yates scale, 208